UPDATING OPTIMAL DECISIONS USING GAME THEORY
AND EXPLORING RISK BEHAVIOR
THROUGH RESPONSE SURFACE METHODOLOGY

THESIS

Jeremy D. Jordan, Captain, United States Air Force

AFIT/GOR/ENS/07-13

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GOR/ENS/07-13

Updating Optimal Decisions Using Game Theory and Exploring Risk Behavior Through Response Surface Methodology

THESIS

Presented to the Faculty

Department of Operational Research

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

Jeremy D. Jordan, B.A. Mathematics

Captain, United States Air Force

March 2007

Updating Optimal Decisions Using Game Theory and Exploring Risk Behavior Through Response Surface Methodology

Jeremy D. Jordan, B.A. Mathematics

Captain, United States Air Force

Approved:

_____   _____

Dr. Marcus Perry, PhD (Co-Chairman)   date

_____   _____

Dr. Sharif Melouk, PhD (Co-Chairman)   date

AFIT/GOR/ENS/07-13

Abstract

This thesis utilizes game theory within a framework for updating optimal decisions based on new information as it becomes available. Methodology is developed that allows a decision maker to change his perceived optimal policy based on available knowledge of the opponents strategy, where the opponent is a rational decision maker or a random component nature. Utility theory is applied to account for the different risk preferences of the decision makers. Furthermore, response surface methodology is used to explore good risk strategies for the decision maker to approach each situation with. The techniques are applied to a combat scenario, a football game, and a terrorist resource allocation problem, providing a decision maker with the best possible strategy given the information available to him. The results are intuitive and exemplify the benefits of using the methods.

Acknowledgements

Table of Contents

List of Figures

List of Tables

Updating Optimal Decisions Using Game Theory and Exploring Risk Behavior Through Response Surface Methodology

I. Introduction

1.1 Background

Updating optimal decisions based on new information as it becomes available is widely applicable in such areas of research as combat scenarios, sports, financial situations, and economic behavior to name a few. In each of these areas, as new information becomes available regarding the possible actions of the opposition or the possible states of nature, the idea of updating an optimal decision policy based on this perception becomes of interest. Optimizing behavior in these situations is of prime interest to military leaders, sports teams, and financial experts who face these decisions. Further exploring the implications of risk behavior in approaching these situations is of great importance as well. In addition, capturing the difference between a perceived optimal strategy and the true optimal strategy will provide insight into the quality of the information perceived. Furthermore, the methodology used to represent rational decision making in the presence of uncertainty in many simulation models is trivialized and an adequate method needs to be developed for general use in these models. The perceived optimal strategy is many times counted as the true optimal strategy. This leads to inaccurate output summary statistics.

1.2 Research Motivation

Bayesian updating is often used solely to update optimal decisions. This method does not consider the action sets and knowledge of what actions are available to either nature or another decision maker. This research considers availability of actions and provides methods to optimize decisions based on this information, as well as

1

techniques for exploring good risk behavior in each situation. Consider the simple combat scenario where a tank observes an unknown object in the distance and must decide what action to take. If the sensors of the tank indicate that the object is an enemy tank, enemy armored personnel carrier, or a friendly tank, the tank may have to choose whether to shoot at the object or investigate the situation further. Depending on the mission, the tank may choose to do one or the other. If the sensors of the tank give updated information that the object is an enemy tank or an enemy armored personnel carrier, the optimal decision of the tank may be to shoot at the object. However, if the true identity of the object is of a friendly nature, the perceived optimal decision to shoot differs from the true optimal decision which may be to advance. The difference between these two decisions can be thought of as the regret of the decision maker.

Consider a second example of an engagement between a red tank and a blue tank. If the sensors of the red tank indicate that the damage level of the blue tank is a mobility/firepower kill, the perceived optimal decision of the red tank may be to move forward and capture troops. However, if the true damage level of the blue tank was only a mobility kill, the perceived optimal decision of the red tank to move forward is less than optimal. The red tank will actually put itself unnecessarily in harms way. Its true optimal decision may be to actually shoot again. This difference between the true and perceived optimal decisions needs to be accounted for during actual battle or in a simulation model.

If the information received and known by a decision maker is imperfect or incomplete, decisions will be affected in proportion to the quality and quantity of the information received. If two entities are engaged in a battle, this information will have some effect on the outcome of this battle.

1.3 Research Objectives

This research aims to address the updating of optimal decisions based on information available by developing methodology that will be useful during the decision making process or as implemented in a simulation model.

The objectives of this research are:

1. Develop a methodology that automatically updates an optimal decision over time based on the information available to a decision maker at each time step

2. Develop the methodology to capture the effects of incomplete or inaccurate information on optimal decision policies by measuring the difference between the perceived optimal decision, that is based on this imperfect information, and the true optimal decision which is based on perfect information

3. Present a technique to explore the implications of decision maker risk behavior and subsequently suggest better alternatives

1.4 Research Approach

Since many decisions are based on the actions of another entity or group and/or the outcome of nature, this research proposes the use of game/decision theoretic concepts to model the decision making process. Game theory is also used to update the optimal decision based on the information available. This research focuses on the modeling of two-player games between entities and two-player games between an entity and nature. An entity may be a team or group of some sort depending on who is making the decisions. A typical situation would include two tanks engaged in combat, with an input from sensors as to the conditions of the vehicles and the overall surrounding operating environment. In the nature case, one tank would make decisions based on what its sensors, measuring some aspect of nature, are reporting to him. This research unravels these problems and suggests an improved modeling approach that could be employed during the decision making process, or implemented in a simulation model, using game theoretic concepts. It focuses on the set of possible

outcomes of the engagements by using the information available to update its optimal strategy. To accomplish this, the game will be written in extensive form to capture all of the possible outcomes. These situations provide the framework from which to collect the necessary data or to distribute questionnaires for subject matter experts. This gives way to the payoff matrices necessary to compute optimal strategies for the players.

If perfect information is available, a perfect decision can be made based on that information. If the information received is less than perfect, the effects of this imperfect information on the outcome of the scenario need to be determined. This is done through measuring the difference between the true optimal decision based on perfect information and the perceived optimal decision possibly based on less than perfect information. The outcome of a scenario given one has perfect information is the optimal outcome. The outcome of a scenario, given one has received less than perfect information, is referred to as the perceived optimal decision. The difference between these two values is the degree of loss incurred because of this bad information. Similarly, this difference between the perceived optimal decision and the true optimal decision can also be thought of as the value of perfect information.

This type of rigid modeling assumes each decision maker approaches the situation in a similar fashion. This is not the case, thus utility theory is used to capture the individual preferences of the decision makers. After the initial game scenario is constructed, any decision maker can be represented through determining his utility of the reward matrix. This provides flexibility to model most situations that may arise.

Additionally, a decision maker, unsure of the risk behavior with which to approach a situation or who suspects his past risk behavior has resulted in less than desirable effects, would benefit from a study on the effects of risk behavior. Response surface methodology is used to explore the effects of risk behavior on the outcome of the games.

4

1.5 Assumptions

The methodology presented herein is dependent on several assumptions, further research could surely be performed to account for most if not all of the assumptions made in this effort.

1. Minimax/Maximin Principle - The players of the game are rational decision makers. Player one is trying to maximize his minimum gain while player two is trying to minimize the maximum gain of player one.

2. Zero-sum - The rewards of the outcome sum to zero. That is, the gain of player one is the same as the loss of player two.

3. Sequential and Simultaneous - This theory is sequential in that each player makes decisions based on one's perception of the available actions to the other player. However, it is simultaneous in that each player makes a decision without knowing the moves of the other player with certainty before the game is played.

4. Non-Cooperative - The players of the game are in a conflict with one another and the chance for cooperative bargaining to arise is zero.

5. Static Rewards - The rewards of the players of the game do not change over time.

6. Compete Information - Each player knows the reward matrix and the initial actions available to the other players with certainty.

1.6 Organization

This thesis is composed of five chapters. Chapter I presents the problem, reasons to improve, and the research method used to solve the problem. Chapter II reviews the literature on game theory and its different applications as well as other methods of updating decisions. Chapter III elucidates the methodology used to approach the problem and gives examples of how to apply the techniques. Chapter IV utilizes this methodology on a combat situation, a football game, and a terrorist resource

allocation problem, showing the effectiveness of the methodology. Chapter V provides conclusions and implications of this research and supplies future direction on the broad possibilities for follow on work to this research.

II. Literature Review

2.1 Introduction

Modeling the updating of optimal decisions and measuring the difference between optimal and perceived optimal decisions during decision processes and in simulation models has received minimal research attention. Game theory has not been used to date as a methodology in a simulation model or during a decision process to update optimal decisions given the information available at the time, nor as a way to measure the difference between a perceived optimal decision and a true optimal decision.

Game theory has been used in many facets since John von Neuman and Oskar Morgenstern published Theory of Games and Economic Behavior in 1950, the first formal application of game theory. John Nash, the father of non-cooperative game theory generalized game theory into different types of games so that it could be used in various venues.

This chapter will review the game theory concepts that are used in this research as well as scour the literature on the origins of some of these concepts. Further, the use of game theory in a military concept will be explored with an emphasis on modeling combat situations. The past attempts of updating optimal decisions using the Bayesian approach are explored as well. The chapter focus is on game theory principles as well as the many uses for game theory in military conflicts.

2.2 Game Theory Concepts/Terminology

Scholars are constantly adding new modifications to the theory of games, allowing it to be used across a broad spectrum of disciplines. This research will utilize the basic concepts of game theory, applying it in a unique manner. To understand the past literature and the ideas presented herein, the basic ideas of game theory must be presented.

Decision theory differs from game theory in that one of the players is a non-rational entity who acts randomly. These games are referred to as games against nature. Decision theory can be thought of as a specific application of game theory.

The remainder of this section involves common game theoretic terms and concepts referenced from [23]. A full list of game theoretic terms can also be found here.

There are 5 basic elements to a game.

1. The players of the game -How many are there and does nature/chance play a role.

2. A complete description of what moves the players can make, the set of all possible player actions.

3. The information available to the players when choosing their actions.

4. A description of the payoff consequences for each player for every possible combination of the actions chosen by each player.

5. A description of the all the preferences of the players over payoffs.

   The following list provides some generic definitions for the theory of games.

Simultaneous Games All players choose actions simultaneously without the knowledge of the strategy chosen by the others players, that is one must anticipate what the opponent will do right now. These moves may actually happen at different times but the actions of the players will be unknown to each other.

Sequential Games Each player makes decisions following a specific order. The players can observe what decisions the other players made before making their decisions. If the information they receive about the other player is truth, the game is considered to be a game of perfect information.

Perfect Information Each player has all of the true information about the moves of another player at the time of their decision.

Imperfect Information The information received by one player concerning the moves of another player up to that point is to some degree in error.

Non-Cooperative Game The players of the game are in direct conflict with one another and situations for compromise do not arise. Cooperative games can involve bargaining contracts outside the specific context of the game.

Payoff Quantitative amount of reward received at the end of each game.

Utility A function of the payoff which changes the payoff relative to the preferences of each individual decision maker.

Zero-Sum Game The payoff to the players at the end of the game sums to zero. One team is the winner and the other team is the loser.

Strategy The set of moves or actions that a player can make in a game. A strategy must be complete and definitive, it must capture every possible decision a player can make given any possible situation that arises during the game.

Pure Strategy An action that a player will follow in every possible attainable situation during a game.

Mixed Strategy A strategy that consists of a set of actions that will be assigned a probability distribution, or a weight as to how often they will be played.

Minimax Principle A principle of playing a game that when utilized, provides a common type of play. It tells the decision maker to choose the maximum of the minimum outcomes of each decision.

Maximax Principle A principle of playing a game that assumes that the best possible scenario will occur. It tells the decision maker to choose the maximum possible outcome of all the decisions. This approach is considered a risk prone method of play.

Nash Equilibrium A set of strategies in which neither player will benefit by changing his or her strategy. In a mixed strategy game, the expected values of the payoffs must be maximized.

Normal Form A matrix representation of the outcomes that could occur at the intersection of each of the decisions of the players. Thus if two players have 5 moves each, the normal form will be a 25 x 25 matrix of payoffs.

2.3   Game Theory in Literature

Nash [15] builds on Von Neumann and Morgensterns' [28] theory of two-person zero-sum cooperative games, in which players form various coalitions, by formulating a theory about non-cooperative games. This is based on the absence of coalitions and assumes that each participant acts independently, without collaboration or communication with any of the others. He proves that every finite non-cooperative game always has at least one or more equilibrium points assuming the players are rational, this being the players' good strategy or strategies.

O'Neill [16] emphasizes through test results that while the minimax theory produces high variability among decision makers, the overall average frequencies of moves and proportion of wins when using the minimax theory was identical to the actual players moves and wins in his experimental test. This validates the use of game theory as a methodology for updating optimal decisions over time. Robinson [20] shows the validity of the method where each player of a game chooses the best pure strategy against the accumulated mixed strategy of his opponent up to that point in time.

Recently, there is an increased interest in determining the proper ways to defend against terrorist attacks. Harris [10] emphasizes the importance of using mathematical methods to combat terrorism. Specifically, game theory is of particular interest in determining optimal allocation of resources to defend against terrorists. He states that a barrier to applying this is that the utility of the players must be considered. Bier [4] talks about the optimal allocation of resources to defend against terrorist attacks. She develops cost functions with probabilities of attacks based on the amount of money used to defend a particular target. The concept of game theory is used to account for the conflict between the enemy and the protector. She also states there

is a need to apply this as a dynamic problem, updating the proper allocations based on new information. Zhuang and Bier [31] find equilibrium strategies for terrorists and country defenders. This is done in the context of resource allocations for terrorist attacks and natural disasters. Sandler [21] also shows the use of game theory in terrorism conflict. His future research recommendations include the need for multi-period game theoretic analysis of terrorist operations. This research will apply exactly to that area.

Arrow [1] discusses the intricacies of operations research and decision theory. Any given problem is first stated as if it were in closed form so that it can be solved using game/decision theory but it still must represent the larger model of truth. Values must be assigned to the physical outcome of the decision situations. These value are really just an estimate of the true value yet these value produce answers that will have implications for future decisions.

Charnes et al [6] study chance constrained games where the players are not fully in control of their strategies. At each time step, random perturbations with known distributions are applied to modify the players' strategy. The selection of the strategies by the players are made before these random variables are formed. Stennek's [25] research verifies the concept of the attraction principle. In contrast with a strictly dominated action which is immediately discarded from use, the attraction principle states that if this strictly dominated action is left in the action space, the action which dominates it should be played with higher probabilities. The research confirms the use of this principle in psychological experiments.

Lipovetsky [13] shows the usefulness of game theory in economic situations and advertising research using zero and non-zero sum games with and without complete information, and cooperative and non-cooperative game theory. The approaches presented are efficient and may broaden applications research in economics.

Game theory has proven itself as a useful tool in many disciplines. It is used in economics to represent economic situations, as well as to describe how actual human

populations behave under different scenarios. Game theory is also used in economics as a normative tool that will suggest how people ought to behave. Game theory has been used in Biology to explain the evolution of the 1:1 sex ratio, explain emergence of communication between animals, and to analyze animal fighting behavior and territoriality. In political science, game theory has been used in such areas as public choice, political economy, and social choice theory, the players often being voters, states, interest groups, and politicians.

2.4 Game Theory in Military Context

This research aims to model combat situations using game theory. The first part of this section gives some general guidance regarding the usefulness of game theory in a military context. Game theory has long been used to model war games; however, not in the context of updating optimal decisions based on the information available about enemy actions. The second part of this section shows some of these direct applications. Although, the idea of updating optimal decisions using game theory has not been used yet, it appears to have tremendous potential within the combat arena. In essence, the military context provides the ideal environment for the application of this methodology.

2.4.1 General Combat Guidance. Whittaker [29] discusses the changing nature of the art of warfare including the shortfalls in our current wargame technology. Specifically, the author argues that game theory provides a flexible and promising framework to model representative strategies for improved automation of behaviors in simulations. The basic concepts of game theory are covered as well as four areas needing expansion for the realization of game theoretic wargaming. Thomas and Deemer [27] express the validity of using games to model combat situations. In their paper, they attack operational gaming as a valid technique for providing a solution to a combat scenario. The uses of game theory in combat situations are given, with an emphasis on the appreciation of what a game solution requires. Athans [2] states

that command and control decision making in simulation models could be advanced significantly via control sciences such as game theory. Pugh and Mayberry [19] delve into the theory of measures of effectiveness for military forces. Although military conflict appears at first glance to be a non-zero sum game, the approximation by a zero-sum game approximates actual combat most effectively. Using a zero-sum payoff function is rationalized to compare alternative combat strategies. They also stress the importance of the measures of effectiveness on the validity of the strategies that are available to each force.

Attrition modeling is a popular combat tool explored by the battle community, with the mathematical analysis developed by Lanchester in 1916. An important study of this using game theory was completed by Cruz et al [9] in which a military air operation was handled using concepts from non-zero sum dynamic game theory. The dynamic nature is achieved through observing the actions of the two forces over time. Solan and Yariv [24] look at a 2-player game where information, perfect or less than perfect, can be purchased by player one regarding the likelihoods of the future actions of player two. The study involves a one-shot game where the payoff of player two is his payoff in the original game, and the payoff of player one is the difference between his payoff in the original game given his information and the cost of the information device he purchased. This concept can be applied to any discipline where one player is interested in private information about the other player in an incomplete information game.

2.4.2 Direct Applications. Berkovitz and Dresher [3] utilize game theory in the analysis of an air war at the tactical level. The strategies in the two-person game are allocation decisions of aircraft among various theater air tasks that maximize the payoff possible of that theater mission. The game is simplified from its original form and major assumptions include that each side is aware of the number of planes the other side holds. Caywood and Thomas [5] apply game theory to the battle of a fighter aircraft and a bomber aircraft. They embrace the idea that the theory of

games provides a valuable framework to evaluate future weapons systems. In their example, the two planes are assumed to be rational decision makers. They also fixed the factors in the engagement, including type, speed, altitude, and flight paths of the aircraft involved. The successful use of game theory also required the payoff function to be common to both participants, thus a zero-sum game. Sujit and Ghose [26] use a game theoretical framework to optimize Unmanned Air Vehicle search routes. At each search step, there is uncertainty in the decisions of the surrounding players and constraints on the flight times of the UAV's that drive the optimal search region.

Perry and Moffat [18] attempt to link battlefield intelligence and measures of combat outcomes together by developing a measure of the knowledge possessed by command and control when making decisions. The paper applies game theoretic concepts to a radar and measures the effects of improved intelligence on combat outcomes. This models a battle composed of many entities, different than a one-on-one engagement. McEneaney et al [14] apply game theory to problems in Command and Control for UAV operations, essentially a game of two forces competing with ground vehicles and UAV's attempting to win by accomplishing a mission. In their situation, one player has perfect information while another player has imperfect information. New theory is developed to deal with this situation, instead of using the standard method. The authors reason that this approach is more applicable to this case, however the computational costs are much greater leading to less fidelity in the model than with the standard method, thus the trade-offs need to be studied prior to implementing this method.

Ozdemirel and Kandiller [17] propose a semi-dynamic model to model land combat at the tactical level. They model individual battles and stages together to compose a game-theoretic setting at combat levels between brigade and platoon, but not at the engagement level. Cruz et al [8] presents a dynamic state-space attrition-type model of a complex military operation involving two opposing forces that can be used to investigate the effectiveness of various game theoretic control strategies applied to a complex system in an intelligent hostile environment. Krichman et al [12]

uses game theory to assist forces with the allocation of resources. The state space is discretized to allow the optimal allocation strategies and value of the game to be calculated. Dynamic programming is used to solve these because the game is continually changing.

2.5 Updating Decisions

Cooman [7] talks about updating beliefs based on incomplete information, however only in the context of Bayesian updating and missing information. A technique is presented that allows the decision maker to account for missing data and incomplete information when calculating probabilities. Sandoy [22] explores alternative Bayesian updating approaches in the application area of a drilling operation. This type of updating requires a procedure that automatically updates an assessment as new information arrives. Several alternatives are presented and the use of Bayes theorem produces a fast and automatic updating procedure. Also noted is the fact that in many cases, the reception of new information opens the need to rethink the entire process and model. This research will address this exact issue.

In conclusion, the idea of updating optimal decisions based on the actions available to other decision makers has not been used anywhere, let alone in combat situations. However, the immediate impact this theory could have on combat situations is apparent and the results could be extremely beneficial to the soldiers in Iraq. By updating the optimal decision as new information is received about the opposition, the soldier can guarantee that he is attacking each situation with the best possible strategy. This will inevitably lead to the preservation of additional lives.

2.6 Conclusion

This chapter shows the broad range of areas game theory is used, with a focus on the military applications. The basics of game theory were provided to introduce some of the general concepts that are used in this thesis. The previous methods for updating optimal decisions are also presented. Although there has been much research

in the field of game theory and decision updating, the methods herein for updating optimal decisions when new information is available have not been developed.

III. Methodology

3.1 Introduction

Updating optimal decisions based on the information available at a given time is trivialized in many simulation scenarios as well as during decision making processes. Updating the optimal decision using game theory is an efficient technique that is widely applicable. The difference between a perceived optimal decision and the true optimal decision is not given proper attention in simulation models as the perceived optimal decision is sometimes counted as the true optimal decision. In the same sense, if uncertainty is present while making a decision, capturing the optimal decision based on perception of the situation is of value. The effects of these perceived optimal decisions impact the outcomes of situations, thus this difference must be accounted for.

This chapter first presents a method to collect the pertinent data needed to accurately model decision making processes or scenarios in a simulation model, and the procedure for setting these up using game theoretic techniques. Secondly, a decision-making methodology is developed that allows for the updating of optimal decisions as a function of information available. Subsequently, a method to account for the absence or malignancy of perceived information is proposed, this being a measure between the true optimal and perceived optimal decision of a situation. Next, a method of accounting for the risk behavior of the individuals involved in the decision making process is presented. Also developed is a method to explore risk behavior through the use of design of experiments and response surface methodology. The analysis is based on the specific situation and the desired outcome of the game subject to the amount of variation willing to be accepted. Finally, example calculations are provided to exemplify how the proposed methods are properly applied.

3.2 Data Collection

Initially, data must be collected to satisfy the desired metric that best represents the situation. For instance, in a financial situation where one is deciding the amount

of disposable income to dedicate to a portfolio consisting of stocks and bonds, an obvious metric to use is currency. In a football game where specific plays are being modeled, the best metric to use might be yards. In some situations there may be numerous metrics that interact with one another. It may be best in these situations to consider a scaled ranking of some sort. For instance, in a combat scenario, the outcomes could be ranked using a Likert scale from -5 to 5, with -5 being the worst possible outcome and 5 the best possible outcome to the situation.

3.2.1 Data Sources.   The data sources that are used to compile the information must be accurate and dependable. For example, to develop the appropriate data so a combat game between two tanks can be properly modeled, it is essential to survey veterans of live combat as subject matter experts (SME). The SME should be knowledgeable in the vehicle or position they are describing. However, this does not necessarily preclude experts who have not actually had combat experience. Their experience in determining proportional differences between the outcomes of scenarios is of primary importance. In general, data should be gathered from an accredited statistical database. This ensures that the results formulated through the use of these techniques are accurate. The insight gained from these techniques is entirely dependent on the quality of the input data.

3.2.2 Components of the Game.   In order to model a decision making process, all components of the game must be well defined.

Player An entity who is playing the game i.e. two football teams or a vehicle such as a tank in combat scenario.

Action Any move that a player can accomplish during a game.

Strategy The set of actions that a player can make in a game. This must be complete and definitive, capturing every possible action a player can make during the game. The strategy is also referred to as an action set.

Reward The quantitative amount of payoff received or lost by each player at the end of each play of a game. The two players payoffs will add to zero during a zero-sum game.

Reward Matrix The matrix of rewards or payoffs associated with each combination of actions.

Initially, the field expert must brainstorm all of the possible actions that could possibly be performed during a game, this set of actions is denoted $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$. The possible actions of the other player must be brainstormed as well, which is denoted $\boldsymbol{\beta} = \{\beta_1, \beta_2, \ldots, \beta_n\}$. From here, all possible combinations of the action sets must be expounded which results in the number of player one and player two combinations, $\kappa = m*n$. Each combination must be assigned a value, all κ combination values make up the reward matrix $\mathbf{R}$. Consider a simple combat scenario where each player is a tank from opposing sides. The scenario is set up such that $\boldsymbol{\alpha} = \{Advance, Retreat\}$ and $\boldsymbol{\beta} = \{Advance, Retreat, Shoot\}$ for the two tanks. The combinations of these actions are ranked on a Likert scale between -5 and 5, with 5 being the best possible outcome. See Table 1 for an example of data collection for this simple combat game. Before the proper data is collected, the players, their strategies, and the associated reward matrix must be well-defined. This is done through conversing with known field experts familiar with allowable assumptions. Now, the decision making process can be setup and solved using game theory.

Table 1:   Sample SME Survey

Tank Actions	Opponent Tank Actions	Rank
Advance	Advance	-2
Advance	Retreat	2
Advance	Shoot	-5
Retreat	Advance	0
Retreat	Retreat	3
Retreat	Shoot	4

3.3 Game Theoretic Setup

To demonstrate the formulation and solution to a game theoretic problem, some new terminology must be introduced from reference [23].

Normal Form Each row or column represents an action and each box represents the payoffs to each player for every combination of actions. Generally, such games are solved using the concept of a Nash equilibrium.

Nash Equilibrium A Nash equilibrium, named after John Nash, is a set of strategies, one for each player, such that no player has incentive to unilaterally change his actions. Players are in equilibrium if a change in strategies by any one of them would lead that player to earn less than if she remained with her current strategy. For games in which players randomize (mixed strategies), the expected or average payoff must be at least as large as that obtainable by any other strategy.

Pure Strategy A single action that a player will follow in every possible attainable situation during a game.

Mixed Strategy A strategy that consists of a subset of actions that will be assigned a probability distribution, or a weight as to how often they will be played. This is a probability of choosing a particular action at some play of the game, these probabilities must sum to 1 over the set of actions.

A thorough overview of game theory can be found in reference [30]. The next few sections present only the necessary ideas for this research.

Each player chooses his best strategy assuming that his opponent knows the best strategy for him to follow, thus he maximizes his minimum gain. His opponent chooses the strategy that allows the other to gain the least, attempting to minimize the other players maximum gain. This assumption is fundamental to the theory of games and is called the minimax/maximin principle. Player one will maximize his

minimum gain while player two will minimize the maximum gain of player one. This assumes also that each player knows the strategies available to the other players.

Initially, the game must be examined to determine if a pure strategy emerges for each player, this is referred to as a saddle point. For a reward matrix $\mathbf{R}$, the condition that must hold for a saddle point to exist is:

$$max(rowminimum) = min(columnmaximum).$$

If a saddle point is not present, the game must be set up as a linear program and solved using the simplex algorithm to determine the optimal mixed strategy for each player of the game. For a reward matrix $\mathbf{R}$ where

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1,n-1} & r_{1,n} \\ r_{21} & r_{22} & & r_{2,n-1} & r_{2,n} \\ \vdots & & \ddots & & \vdots \\ r_{m-1,1} & r_{m-1,2} & & r_{m-1,n-1} & r_{m-1,n} \\ r_{m1} & r_{m2} & \cdots & r_{m,n-1} & r_{m,n} \end{bmatrix}$$

and the normal form of a game given in Table 2,

Table 2:   Normal Form of a Game

	β_1	β_2	$\cdots$	β_{n-1}	β_n
α_1	r_{11}	r_{12}	$\cdots$	$r_{1,n-1}$	$r_{1,n}$
α_2	r_{21}	r_{22}		$r_{2,n-1}$	$r_{2,n}$
$\vdots$	$\vdots$		$\ddots$		$\vdots$
α_{m-1}	$r_{m-1,1}$	$r_{m-1,2}$		$r_{m-1,n-1}$	$r_{m-1,n}$
α_m	r_{m1}	r_{m2}	$\cdots$	$r_{m,n-1}$	$r_{m,n}$

the following linear program can be setup

$$\max z = \nu + 0w_1 + 0w_2 \ldots + 0w_m$$

$$s.t. \, \nu \le r_{11}w_1 + r_{21}w_2 + \cdots + r_{m1}w_m$$

$$\nu \le r_{12}w_1 + r_{22}w_2 + \cdots + r_{m2}w_m$$

$$\vdots$$

$$\nu \le r_{1n}w_1 + r_{2n}w_2 + \cdots + r_{mn}w_m$$

$$\sum_i w_i = 1; w_i \ge 0; \forall i = 1, \ldots, m$$

$$(1)$$

where $\gamma = \{w_1, w_2, \ldots, w_m\}$ denotes a probability distribution assigned to the set of actions $\boldsymbol{\alpha}$, the strategy of player one. ν is the value of the game, which player one is maximizing in the objective function over the actions of player two, accounted for in the constraints. This linear program can be easily solved via the simplex method to solve for the values of this strategy. To compute the strategy $\delta = \{\delta_1, \delta_2, \ldots, \delta_n\}$ of the action set $\boldsymbol{\beta}$ of player two, the dual linear program of Equation (1) must be formulated. This is easily set up.

$$\min z = \omega + 0\delta_1 + 0\delta_2 \ldots + 0\delta_n$$

$$s.t. \, \omega \ge r_{11}\delta_1 + r_{12}\delta_1 + \cdots + r_{1n}\delta_n$$

$$\omega \ge r_{21}\delta_2 + r_{22}\delta_2 + \cdots + r_{2n}\delta_n$$

$$\vdots$$

$$\omega \ge r_{m1}\delta_1 + r_{m2}\delta_2 + \cdots + r_{mn}\delta_n$$

$$\sum_j \delta_i = 1; \delta_j \ge 0; \forall j = 1, \ldots, n$$

$$(2)$$

Formally stated, a pure strategy occurs when the strategy for the action sets $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$ have the following properties: $w_i = 1$ for 1 of m actions and $w_i = 0$ $\forall$ remaining i's or $\delta_j = 1$ for 1 of δ actions and $\delta_j = 0$ $\forall$ remaining j's, respectively. A mixed strategy occurs when the strategy for the action sets $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ consists of a probability distribution for the actions that will be played, γ and δ respectively.

3.3.1 Value of the Game. The players' mixed strategies result in a floor value for player one denoted π. In other words, player one is guaranteed to receive no less than π if his mixed strategy $\gamma = \{w_1, w_2, \ldots, w_m\}$ is played. π is also the ceiling value for player two, guaranteeing him from losing any more than π by playing his mixed strategy $\delta = \{\delta_1, \delta_2, \ldots, \delta_n\}$. When the value of the game to each of the players is equal, a Nash equilibrium occurs. Consequently, any mixed strategy that results in equal values of π meets this criteria and is considered an optimal strategy. It is important to note that π is the expected value to each of the players over time. At each play of a game, there will be some variation from π. As time approaches infinity however, each player can expect their reward to approach this value, π.

The common value of the game π for each player is actually the solution to their respective linear programs, $\pi = \nu = \omega$, however since the reward matrix $\mathbf{R}$ will be manipulated in upcoming sections to account for player preferences, the equation

$$
\begin{aligned}
\pi &= [\gamma = \{w_1, w_2, \ldots, w_m\}] * \mathbf{R} * [\delta = \{\delta_1, \delta_2, \ldots, \delta_n\}]' \\
&= \gamma \, \mathbf{R} \, \delta'
\end{aligned}
\tag{3}
$$

will be used herein. Note, the value of δ used for nature in the one-player versus nature case is a uniform distribution across the strategy of nature. This is a valid assumption because when information is unavailable regarding the likelihood of nature, the uniform distribution provides the best estimate. In the two-player game, δ will be the actual strategy used by player two according to the minimax principle.

3.4 Updating the Optimal Decision

The optimal strategy will change depending on the data available about the action sets or strategy of the players of the game. The strategy γ for the action set $\boldsymbol{\alpha}$ at time step s is dependent on a player's perception of what actions are available to the other player. That is,

$$\hat{\gamma}^{(0)} = [\gamma = \{w_1, w_2, \ldots, w_m\} | \hat{\beta}^{(0)} = \{\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_n\}]$$

for $s = 0$ at the start of the game. This shows that the strategy of player one, γ, is based on his perception of what actions are available to player two, which is all of the possible actions of player two initially. In general,

$$
\begin{aligned}
\hat{\gamma}^{(s)} &= [\gamma = \{w_1, w_2, \ldots, w_m\} | \hat{\beta}^{(s)} = \{\beta_1, \beta_2, \ldots, \beta_o\}] \\
&= \gamma | \hat{\beta}^{(s)}
\end{aligned}
\tag{4}
$$

where o is the number of actions perceived by player one. The inequality $o \leq n$ holds, implying that player one can only perceive as many actions as he originally knows player two is capable of choosing from. This will continue up to time step q, the number of time steps in the game. As gamma updates, it is dependent on data obtained from some source or combination of sources that is perceiving the situation, or information about $\boldsymbol{\beta}$. Thus $\hat{\boldsymbol{\beta}}$ is dependent on

$$\boldsymbol{\zeta} = \{\zeta_1 \cap \zeta_2 \cap \ldots \cap \zeta_k\},$$

where ζ_i is source i of k number of sources. Thus,

$$
\begin{aligned}
\hat{\beta}^{(s)} &= [\beta | \boldsymbol{\zeta}^{(s)} = \{\zeta_1 \cap \zeta_2 \cap \ldots \cap \zeta_k\}] \\
&= \beta | \boldsymbol{\zeta}^{(s)}
\end{aligned}
\tag{5}
$$

where $\boldsymbol{\zeta}^{(s)}$ is the set of sources available at time s. Thus,

$$\hat{\gamma}^{(s)} = [\gamma | \{\hat{\beta}^{(s)} | \boldsymbol{\zeta}^{(s)}\}] \tag{6}$$

showing that the strategy of player one is dependent on the information received from the sensor and his perception of the action set of player two.

The reward matrix $\mathbf{R}$ and strategy of player two δ will update at each time step as well depending on the player's perceptions of the action sets available to the other players, denoted $\hat{R}^{(s)}$ and $\hat{\delta}^{(s)}$, respectively.

3.4.1 Difference Between Optimal and Perceived Optimal Decisions. During a game, a player may not have true information about his opponents set of available actions. The mixed strategy may not be the proper mixed strategy to use since it may be based on false information, making it a perceived optimal mixed strategy, denoted above as $\hat{\gamma}^{(s)}$. The true optimal mixed strategy based on perfect information may be used to determine how poor this perceived mixed strategy is. Before introducing this technique, it is necessary to pioneer some fresh terminology from reference [23].

Sequential game These games occur when each player makes decisions following a specific order. The players can observe what decisions the other players made before making their decisions. If the information they receive about the other player is truth, the game is considered to be a game of perfect information.

Perfect Information Occurs in a game when one player has all of the true information about the actions of the other players or possible sets of actions at the time of their decision.

Imperfect Information Occurs in a game when the information received by one player concerning the actions of the other players or possible sets of actions up to that point is to some degree in error.

The true optimal strategy is thus denoted $\gamma^{(s)}$. Similarly, the true reward matrix is $\mathbf{R}^{(s)}$ and the true strategy of player two is $\delta^{(s)}$. If the information received via

sensors or some other source $\boldsymbol{\zeta} = \{\zeta_1, \zeta_2, \ldots, \zeta_k\}$ is less than perfect (imperfect), a difference will occur between the optimal strategy and the perceived optimal strategy. That is, $\hat{\gamma}^{(s)}$ is less than or equal in quality to $\gamma^{(s)}$. The magnitude of the difference between the perceived optimal and true optimal strategies can be measured using the value of the game. The value of the game π provides the best comparison measure for contrasting two strategies. The value corresponding to the perceived optimal strategy $\hat{\gamma}^{(s)}$ is calculated using the perceived optimal strategy, the true reward matrix, and the true strategy of player two,

$$\hat{\pi}^{(s)} = \hat{\gamma}^{(s)} \, \mathbf{R}^{(\mathbf{s})} \, \delta^{(\mathbf{s})'}. \tag{7}$$

The value of the true optimal strategy is similarly calculated

$$\pi^{(s)} = \gamma^{(s)} \, \mathbf{R}^{(\mathbf{s})} \, \delta^{(\mathbf{s})'}. \tag{8}$$

The difference between the values of the two strategies is

$$\vec{\pi}^{(s)} = \pi^{(s)} - \hat{\pi}^{(s)}. \tag{9}$$

This value, $\vec{\pi}^{(s)}$, will change as the game progresses. Initially the difference between the value of the game of the perceived and the true optimal decision is $\vec{\pi}^{(0)} = 0$. The value, $\vec{\pi}^{(s)}$, can be used to determine the value of obtaining perfect information and also to measure the value of the information obtained from sources $\boldsymbol{\zeta} = \zeta_1, \zeta_2, \ldots, \zeta_k$. Note, this value may or may not be representative of the actual value because of the nature of the measure. For example, in a ranking system, $\vec{\pi}^{(s)}$ will provide a frame of reference for which two different decisions can be compared and/or the value of the source can be observed over time.

3.5 Determining User Preferences

Traditional zero-sum game theory assumes that each player will approach the game in an identical fashion, however this is not always the case. While it is true that each player will attempt to maximize his minimum gain and minimize the other players maximum gain, the values of the reward matrix $\mathbf{R}$ will not always reflect the true value of the reward to each player. This difference in value is accounted for by determining the players' risk preference and changing the values in the reward matrix to reflect this preference.

3.5.1 Utility. Since the reward matrices are, in essence, just data of whatever measure is decided upon to represent the game, utility theory must be used to represent what value the reward of the situation has to each player. That is, the true values of the situations to the players in some cases, when all other influencing factors are considered, is different than the general reward matrix. The original reward matrix will produce strategies that can be thought of as the expected case, the strategy a normal rational decision maker would take in a situation. Furthermore, utility will allow every type of decision maker to be represented based on their individual risk taking behavior. That is, a decision maker may be risk averse, risk neutral, or risk prone, of which there are different levels. The risk averse individual will avoid risk more so than the risk neutral individual in the expected case, going for the sure thing. The risk prone individual will approach situations with great risk in comparison with the risk neutral individual; he will attempt to maximize his payoff regardless of the chance for loss. The risk neutral individual will approach the situation as the average rational individual would, maximizing his minimum payoff for the given reward matrix. This is explained better through an inspection of the reward matrix. For a reward matrix $\mathbf{R}$ with strategy $\gamma = \{w_1, w_2, \ldots, w_m\}$ and action set $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, an action α_i is considered a high risk action if its standard deviation σ_i is large in respect to the other α_i's σ_i's. The risk prone individual will play this action more often. An action α_i is considered a low risk action if its standard

deviation σ_i is small in respect to the other α_i's σ_i's. The risk averse individual will use this action more often. This classification of the actions is also dependent on the mean of choosing a particular action across the different possibilities of the strategy of player two. This mean is calculated assuming a uniform distribution as the strategy of player two and multiplying the reward across a particular action by this uniform distribution. For example, an action with a low mean in respect to the other actions and a high standard deviation may not be treated as a risk prone action when the game is solved because the mean is too low to allow it even to be considered as a viable action.

The utility matrix is conjectured using some function of the original payoff and the risk tolerance level, which accounts for the players' risk behavior. A popular, widely used function is the exponential utility function which has been shown to be extremely useful in evaluating risk behavior. The function transforms the original reward matrix using a parameter, ρ, which accounts for the players' risk behavior. There are several forms of the exponential utility function, some may work better for certain cases than others. From [11], for the monotonically increasing measure,

$$u\{R_{ij}\} = \frac{1 - exp[-(R_{ij} - Low)/\rho]}{1 - exp[-(High - Low)/\rho]} \; , \tag{10}$$

where $u\{R_{ij}\}$ is the utility of the $ijth$ element of the reward matrix, Low is the lowest level of the measure m, High is the highest level of m, and ρ is the exponential constant, or risk preference, for the value function. Recall, m is the measure used to compile the original reward matrix and is the basis for computing the expected case strategy. Equation 10 will alter this original measure to produce the true value of the original measure to the decision maker.

Varying levels of the risk tolerance will produce different types of risk behavior, evident in the strategies of the players during game theoretic engagements. The actual approach to determine ρ for each player is addressed in upcoming sections.

3.5.2 Rho Assumptions. As the game updates, the risk behavior of the players will evolve as well. A player may initially approach the game with a risk averse attitude, then transition to a risk prone attitude as the game progresses. In a one player versus nature scenario, only the risk attitude of player one needs to be considered, and the results are just dependent on this risk attitude. Because of the assumptions present during a two-player game, the values of ρ take on a slightly different meaning. Examining the implications of the game theoretic setup from the perspective of player one implies his decisions are based on the assumption that player two knows his risk strategy and is attempting to minimize the maximum gain of player one. This is really the zero-sum assumption. This is also the actual risk attitude used by player two. The value of the game can be calculated for these different risk attitudes,

$$\hat{\pi}_{\rho_1(\rho_2)}^{(s)} = \hat{\gamma}_{\rho_1}^{(s)} \, R^{(s)} \, \delta_{\rho_2}^{(s)'}.$$

See Table 3 for a description of an example of high and low risk tolerance levels and their meanings. In general, ρ approaching zero from infinity indicates a more risk averse behavior while ρ approaching zero from negative infinity indicates a more risk prone behavior. $\rho = 0$ is undefined for the exponential utility function and $\rho = \infty$ is risk neutral behavior.

Table 3:   Example Risk Tolerance Levels

-50	Risk Neutral
-1	Risk Prone
1	Risk Averse
50	Risk Neutral

3.5.2.1 Certainty Equivalent. To actually determine the risk preference parameter ρ_i of player i, for $i = 1, 2$, the certainty equivalent for each player must be gathered. The certainty equivalent is the certain payoff a decision maker will accept to avoid a given gamble. The certainty equivalent is obtained through a lottery presented to the decision maker. The procedure for determining ρ for each player of

the game will be presented in chapter four. The general concept of ρ is presented here so a technique for studying risk behavior can be developed in the next section.

3.5.3 Studying the Effects of Risk Behavior.   The combinations of levels of ρ can be examined through plotting the different risk preferences of the two players versus the value of the game. This can be observed on a three dimensional graph of the surface, a contour plot of the three dimensional surface, or interaction plots of the two variables and the response which is the value of the game, π.

3.5.3.1 Design of Experiments.   Exploring the varying levels of risk tolerance and their effects on the value of the game, π, through a designed experiment is an efficient way to characterize their relationship to one another. If the analysis needs to be done real time, and there is not time to look at every combination of levels of risk tolerance, a design of experiment can be run to examine the high and low levels of the risk tolerances and their effects on the value of the game. An adequate model can be gathered in an efficient manner. This is important as it will provide the value of the game as a function of the risk tolerances of the two players. This function can then be optimized using response surface methodology.

The two factors that are varied are the risk tolerances of the two players, of which the high and low levels will be examined. From Table 3, it is inferred that each factor has 4 levels, discontinuous at 0. Note, the high and low levels of ρ will be dependent on the certainty equivalent and the range of the values in the reward matrix. These high and low values of ρ are for one example game. The exact procedure for determining these levels is presented in chapter four. Three indicator variables are used to aid in setting up a proper design matrix, representing the four different combinations of the risk tolerances of the players. The resulting design matrix X is given in Table 4. The design regions and graphical representation of the design matrix is given in Figure 1.

Table 4:   Design Matrix(Original/Coded)

ρ_1	ρ_2	x_1	x_2	x_3	ρ_1	ρ_2	x_1	x_2	x_3
-50	1	1	0	0	-1	-1	1	0	0
-1	1	1	0	0	1	-1	1	0	0
-50	50	1	0	0	-1	1	1	0	0
-1	50	1	0	0	1	1	1	0	0
-50	-50	0	1	0	-1	-1	0	1	0
-1	-50	0	1	0	1	-1	0	1	0
-50	-1	0	1	0	-1	1	0	1	0
-1	-1	0	1	0	1	1	0	1	0
1	1	0	0	1	-1	-1	0	0	1
50	1	0	0	1	1	-1	0	0	1
1	50	0	0	1	-1	1	0	0	1
50	50	0	0	1	1	1	0	0	1
1	-50	0	0	0	-1	-1	0	0	0
50	-50	0	0	0	1	-1	0	0	0
1	-1	0	0	0	-1	1	0	0	0
50	-1	0	0	0	1	1	0	0	0

The response matrix y, the value of the game π, is calculated by setting the levels of ρ_1 and ρ_2, transforming the reward matrix accordingly, then solving the game for each of the transformed matrices. y is different for each game played. The resulting model is obtained using regression techniques, namely least squares estimation:

$$(X'X)^{-1}X'y. \tag{11}$$

The resulting model is

$$\hat{y} = \theta_0 + \theta_1\rho_1 + \theta_2\rho_2 + \theta_3 x_1 + \theta_4 x_2 + \theta_5 x_3 + \theta_{12}\rho_1\rho_2 + \theta_{13}\rho_1 x_1 + \theta_{14}\rho_1 x_2 +$$

$$\theta_{15}\rho_1 x_3 + \theta_{23}\rho_2 x_1 + \theta_{24}\rho_2 x_2 + \theta_{25}\rho_2 x_3 + \theta_{123}\rho_1\rho_2 x_1 + \theta_{124}\rho_1\rho_2 x_2 + \theta_{125}\rho_1\rho_2 x_3 \tag{12}$$

Since there is no residual error present, the variability in the y matrix is solely due to the changes in risk behavior of the two players, thus only one replication needs to be run. Now, the model can be optimized using robust parameter design. The main objective is to maximize the response subject to some level of variation willing
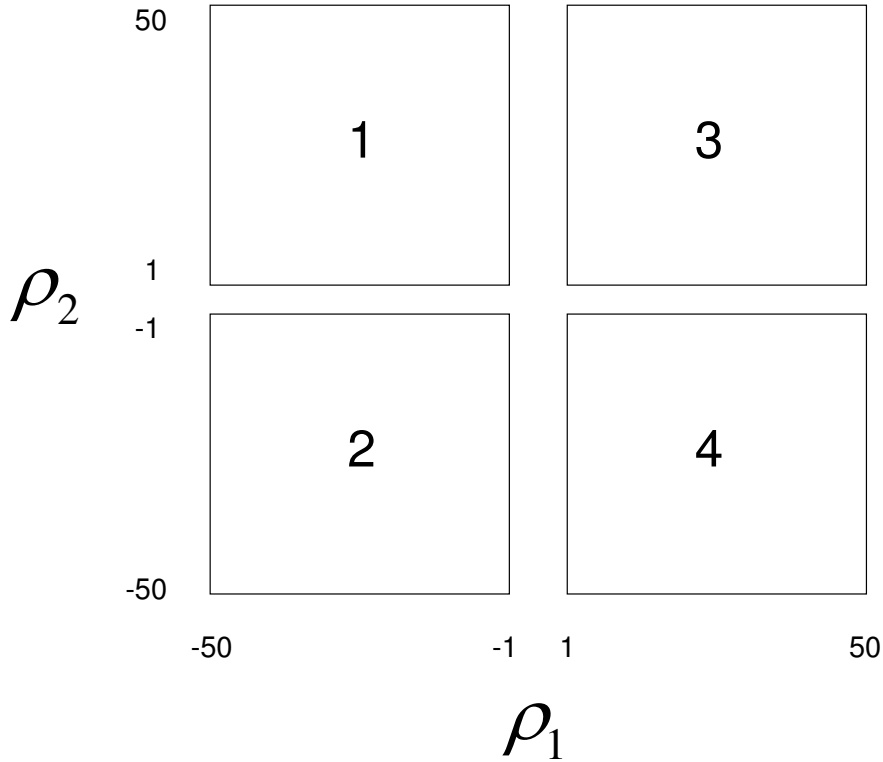
Figure 1:   Design Region

to be accepted. Conversely, the variation can be minimized subject to some constraint on the acceptable level of the response.

3.5.4   Robust Parameter Design.   It is important to note, while absolute optimization will be of value to the players of the game, formulating the response surface as a function of the risk tolerances and learning about the process is of greater importance. Approaching the problem from the perspective of player one sets the risk tolerance of player one as the control factor and the risk tolerance of player two as the noise factor, the uncontrollable factor. We are interested in not only the main effects of the control and noise factors, but also the control $\times$ noise interactions as they describe the variance in the response. In fact, all interactions will describe the variation in the response but the control by noise interactions can be exploited to help

design robust systems. A convenient way to examine the effects of risk behavior on the value of the game is through inspection of the response surface, interaction plots of the response, and contour plots of the estimated surface. These techniques have not been used to explore risk behavior in the past. They will be examined in detail in chapter four.

3.5.5 Optimizing Risk Strategy in the Game Against Nature.   In the case of the game against nature, an optimization problem for risk strategy would be fairly simple. The value of the game π can be maximized across levels of ρ subject to some constraint on the variability in π across levels of ρ. The mean and variance response surface models are only functions of the risk tolerance of player one, subject to the hypothesized distribution of nature. In fact, the variance model is a function of ρ, however it represents variance in y across values of ρ. If the mean response surface was linear and the variance non-linear, the decision maker could set up a linear program by performing a LaGrangian Relaxation on the non-linear constraint. The mean could be maximized subject to a constraint on the amount of variance willing to be accepted. Conversely, a non-linear program could be set up to minimize variance subject to some constraint on the desired mean.

3.6 Example Calculation for the 1-Player vs Nature Game

This research aims to model situations where a player is competing against an outcome of nature. A natural situation where this occurs is on the battlefield during a war. Consider a game where player one is a U.S. Army tank and player two is the data being received from sensors on the battlefield. This section will demonstrate the theory presented above on this scenario.

3.6.1 Modeling a 1-Player vs Nature Combat Game.   Initially, some assumptions of the conditions under which this battle is being conducted must be stated. The battle is set on a normal battleground, i.e. no urban clutter, etc., in the desert of Iraq. The tank is observing an unknown object with its onboard sensors and is engaged

with a hostile enemy whom the tank is attempting to overtake. Civilian casualties are a concern and need to be minimized. Initially, the players action sets for this situation must be determined through surveys of SMEs. The combinations of these actions are then ranked using a Likert scale between -5 and 5. After collecting the necessary data associated with all combinations of these action sets, the normal form of the game is given in Table 5.

Table 5:   Normal Form of Game against Nature

	Enemy Truck	Civilian Truck	Enemy Tank	Friendly Tank
Fire Mortar	4	-4	5	-5
Advance	1	4	0	4
Do Nothing	-1	1	-2	1

The action set of player two is

$$\boldsymbol{\beta} = \{\beta_1, \beta_2, \beta_3, \beta_4\} = [EnemyTruck, CivilianTruck, EnemyTank, FriendlyTank].$$

This is the actual data received by the sensor and is considered nature, or player two. The action set of player one

$$\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \alpha_3\} = [FireMortar, Advance, DoNothing],$$

are the possible actions to take based on the information received from the sensor. In reality, these actions set will be much larger and more complex. Setting the game up to determine the optimal strategy for player one requires solution of the linear

34

program

$$\max z = v + 0w_1 + 0w_2 + 0w_3$$

$$s.t. \ v \le 4w_1 + 1w_2 + -1w_3$$

$$v \le -4w_1 + 4w_2 + 1w_3$$

$$v \le 5w_1 + 0w_2 + -2w_3$$

$$v \le -5w_1 + 4w_2 + 1w_3$$

$$\sum_i w_i = 1$$

$$w_i \ge 0 \ \forall \ i.$$

(13)

Using the Simplex method to solve the linear program yields the mixed strategy $\gamma = \{w_1, w_2, w_3\} = \{.2857, .7143, 0\}$ for player one. Player one shoots his mortar roughly 29% of the time, advances 71% of the time, and should always do something in this situation. The value of the game to player one, assuming uniformity across actions of player two, is

$$\pi \ = \ \gamma \, \mathbf{R} \, \delta'$$

$$= \ [.2857, .7143, 0] \begin{bmatrix} 4 & -4 & 5 & -5 \\ 1 & 4 & 0 & 4 \\ -1 & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} .25 \\ .25 \\ .25 \\ ,25 \end{bmatrix}$$

$$= \ 1.6072.$$

This value is the expected reward that player one will gain on each play of the game. In this case, the rewards are on a scale so a higher value simply means better, where 5 is the best. The uniform distribution is used for the strategy of nature as this provides the best estimate when information is unavailable. With information available about

the actual probabilities from nature, these probabilities could be updated and thus provide a better approach.

   3.6.2   Updating the Optimal Solution.   Suppose the tank (player one) received information from an onboard sensor that fire was detected from the object and was directed at the tank. His optimal decision will now change based on the fact that his perception of the object is that of enemy nature,

$$
\begin{aligned}
\hat{\beta}^{(1)} &= [\beta | \zeta^{(1)} = \{\zeta_1 = OnboardSensor\}] \\
&= [EnemyTruck, EnemyTank].
\end{aligned}
$$

The perception by the tank due to data received by the tanks onboard sensor is that the object is an enemy vehicle. The original optimal strategy of player one

$$
\hat{\gamma}^{(0)} = \{.2857, .7143, 0\}
$$

changes to the updated strategy

$$
\begin{aligned}
\hat{\gamma}^{(1)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{1, 0, 0\}.
\end{aligned}
$$

Knowing that the object is an enemy, player one should shoot his mortar 100% of the time. Suppose the tank received an update from an airborne sensor that the object being observed was possibly a civilian truck. Now,

$$
\begin{aligned}
\hat{\beta}^{(2)} &= [\beta | \zeta^{(2)} = \{\zeta_1 = OnboardSensor \cap \zeta_2 = AirborneSensor\}] \\
&= [EnemyTruck, CivilianTruck, EnemyTank]
\end{aligned}
$$

results in the strategy

$$\hat{\gamma}^{(2)} = [\gamma | \hat{\beta}^{(2)}]$$
$$= \{.3077, .6923, 0\}$$

with $\pi = 1.5385$ being the perceived value of the game. If another sensor reported the object was either a friendly tank or civilian truck, the first sensor may be considered obsolete. Thus,

$$\hat{\beta}^{(3)} = [\beta | \zeta^{(3)} = \{\zeta_2 = AirborneSensor \cap \zeta_3 = OtherSensor\}]$$
$$= [CivilianTruck, FriendlyTank]$$

results in the strategy

$$\hat{\gamma}^{(3)} = [\gamma | \hat{\beta}^{(3)}]$$
$$= \{0, 1, 0\}$$

with $\pi = 4$ being the perceived value of the game.

3.6.3 Difference Between Optimal and Perceived Optimal Strategies. During a post-war analysis, suppose a decision-maker was interested in the performance of a particular sensor during the campaign, take the onboard sensors for instance. The value of the game based on information from the sensor can be calculated and measured against the truth. Recall at $s = 1$,

$$\hat{\beta}^{(1)} = \{EnemyTruck, EnemyTank\}$$

and

$$\hat{\gamma}^{(1)} = \{1, 0, 0\}$$

with $\pi = 4$. From observing the information from $s = 2$ to $s = 3$ and using tapes from the airborne sensors, we assume that the onboard sensors have given a report that is less than accurate. To gauge the quality of this information, first calculate the true values based on truth, obtained from post-war analysis. Assume the true actions available to nature are

$$\beta^{(1)} = [CivilianTruck, FriendlyTank].$$

This produces a true value

$$
\begin{aligned}
\pi^{(1)} &= \gamma^{(1)} R^{(1)} \delta^{(1)'} \\
&= [0, 1, 0]
\begin{bmatrix}
-4 & -5 \\
4 & 4 \\
1 & 1
\end{bmatrix}
\begin{bmatrix}
.5 \\
.5
\end{bmatrix} \\
&= 4,
\end{aligned}
$$

the true value of the game at $s = 1$. The value of the perceived optimal decision is calculated similarly using

$$\hat{\beta}^{(1)} = [EnemyTruck, EnemyTank].$$

with

$$\hat{\gamma}^{(1)} = \{1, 0, 0\}.$$

38

The value of the perceived optimal strategy is

$$
\begin{aligned}
\hat{\pi}^{(1)} &= \hat{\gamma}^{(1)}\, R^{(1)}\, \delta^{(1)'} \\
&= [1,0,0] \begin{bmatrix} -4 & -5 \\ 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} .5 \\ .5 \end{bmatrix} \\
&= -4.4,
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\vec{\pi}^{(1)} &= \hat{\pi}^{(1)} - \pi^{(1)} \\
&= -4.4 - 4 \\
&= -8.4,
\end{aligned}
$$

showing that there was a significant difference in the value of the game between the perceived optimal decision and the true optimal decision. The value lost by player one due to the information he received from the onboard sensors regarding the object was 8.4, an extreme number considering the range of **R** is 10.

3.7 Example Calculations for the 2-Player Game

This section explores an intuitive example of the implementation of the presented methodology on a two-player game.

3.7.1 Setting up a 2-Player Game. This research seeks to model 2-player games between competing entities, where each player is an entity of opposing forces or teams. To demonstrate how to set up a game in this manner, consider the scenario of an NFL football game, the Green Bay Packers (player one) versus the Chicago Bears (player two) at Lambeau Field. The fidelity of the game is at the play-by-play level, offense versus defense. Each play of the game will have a different action set for each player, dependent on the time of the game, score, etc. Let player one

(offense) approach this specific play with the action set $\boldsymbol{\alpha} = \{1, 2, \ldots, m\}$ being $\boldsymbol{\alpha} = [Option, Deep\ Pass, Short\ Pass, Run\ up\ the\ Gut]$ and player two (defense) with the action set $\boldsymbol{\beta} = \{1, 2, \ldots, n\}$ be $\boldsymbol{\beta} = [4-3\ Prevent, 4-3\ Man, 4-4\ Man\ no\ blitz, 5-4\ Man\ blitz]$. $\mathbf{R}$, the reward matrix, is found by estimating the amount of yards gained by the offense (and subsequently lost by the defense) at each of the different combinations of the action sets of the players. This could perhaps be gathered through past statistics of an average number of yards gained in each situation. The normal form of the game is given in Table 6. The maximum of the row minimums is 4 in $R_{3,3}$ corresponding to the short pass. The minimum of the column maximums is 4.4 in $R_{4,3}$ corresponding to the 4-4 Man no blitz. Since $R_{3,3} \neq R_{4,3}$, a saddle point does not exist and the game must be solved using linear programming.

Table 6:   Normal Form of 2-player Game

	4-3 Prevent	4-3 Man	4-4 Man no blitz	5-4 Man blitz
Option	9.4	4.75	3.5	4.9
Deep Pass	2.1	3.75	4.2	15
Short Pass	7.3	4.5	4	5.1
Run up the Gut	4.9	4.1	4.4	3.2

Setting up the linear program to solve the game and extract the optimal strategy γ for the action set $\boldsymbol{\alpha}$ of the first player yields

$$\max z = v + 0w_1 + 0w_2 + 0w_3 + 0w_4$$

$$s.t.\ v \leq 9.4w_1 + 2.1w_2 + 7.3w_3 + 4.9w_4$$

$$v \leq 4.75w_1 + 3.75w_2 + 4.5w_3 + 4.1w_4$$

$$v \leq 3.5w_1 + 4.2w_2 + 4w_3 + 4.4w_4$$

$$v \leq 4.9w_1 + 15w_2 + 5.1w_3 + 3.2w_4$$

$$\sum_i w_i = 1$$

$$w_i \geq 0\ \forall\ i.$$

$$(14)$$

Player one would thus approach the play by choosing an action according to the distribution $\gamma = \{0, .0271, .3801, .5928\}$, or 3 percent of the time player one should choose to call a deep pass, 38 percent of the time player one should choose to call a short pass, and 59 percent of the time player one should choose to run up the gut. Similarly, the optimal strategy δ for the second player can be found by setting up the dual of Equation (14)

$$\min z = \omega + 0\delta_1 + 0\delta_2 + 0\delta_3 + 0\delta_4$$
$$s.t.\ v \geq 9.4\delta_1 + 4.75\delta_2 + 3.5\delta_3 + 4.9\delta_4$$
$$v \geq 2.1\delta_1 + 3.75\delta_2 + 4.2\delta_3 + 15\delta_4$$
$$v \geq 7.3\delta_1 + 4.5\delta_2 + 4\delta_3 + 5.1\delta_4$$
$$v \geq 4.9\delta_1 + 4.1\delta_2 + 4.4\delta_3 + 3.2\delta_4$$
$$\sum_j \delta_j = 1$$
$$\delta_j \geq 0\ \forall\ j.$$

$$(15)$$

Using the simplex method again to solve the linear program yields the mixed strategy $\delta = \{0, .4364, .5415, .0221\}$ for player two. Player two should call the 4-3 Man 44 percent of the time, 4-4 Man no blitz 54 percent of the time, and 5-4 Man blitz 2 percent of the time in this particular situation. These mixed strategies allow for randomness while play calling which keeps the opponent honest. For instance, on 2nd down and 2, there will be some probability of calling a long pass.

A peculiar consequence of these mixed strategies occurs when one particular action appears to get better in terms of the reward matrix. Perhaps the expected yards gained for the option increases in the reward matrix, the play is working better than in past seasons. The Nash equilibrium produces counterintuitive results in that you may actually use the option less than when it previously was producing lower expected yards. This phenomenon is due to the interactive nature of the game, player

two knows that the option is working better (the reward matrix is common knowledge) for player one so he will decide to increase his defense against it. Player one must choose to increase calling a different play in order to prove to the defense that he will not be running the option. This further validates the use of this technique for the game of football as this is often the mindset of the coaches.

3.7.2 Value of the Game. The value of the football game above is

$$
\begin{aligned}
\pi &= \gamma \, \mathbf{R} \, \delta' \\
&= \begin{bmatrix} 0, 0.0271 & 0.3801 & 0.5928 \end{bmatrix}
\begin{bmatrix}
9.4 & 4.75 & 3.5 & 4.9 \\
2.1 & 3.75 & 4.2 & 15 \\
7.3 & 4.5 & 4 & 5.1 \\
4.9 & 4.1 & 4.4 & 3.2
\end{bmatrix}
\begin{bmatrix}
0 \\
0.4364 \\
0.5415 \\
0.0221
\end{bmatrix} \\
&= 4.2425
\end{aligned}
$$

By playing his mixed strategy, player one expects to gain 4.24 yards with each play of the game. Player two, by playing his mixed strategy expects to give up no more than 4.24 yards with each play of the game in the long run.

3.7.3 Updating optimal play calling. If the quarterback approached the line and observed the defense, then concluded that they were not in a 4-3, but possibly a 4-4 or a 5-4, his mixed strategy would change due to this new information. The normal form with the updated reward matrix is shown in Table 7.

Table 7: Updated 2-player Game

	4-4 Man no blitz	5-4 Man blitz
Option	3.5	4.9
Deep Pass	4.2	15
Short Pass	4	5.1
Run up the Gut	4.4	3.2

The perception of the strategy of player two by player one is dependent on the observation of the quarterback,

$$
\begin{aligned}
\hat{\beta}^{(1)} &= [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = QBObservation\}] \\
&= [4 - 4Mannoblitz, 5 - 4Manblitz]
\end{aligned}
$$

The updated strategy is

$$
\begin{aligned}
\hat{\gamma}^{(1)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{0, .1, 0, .9\}.
\end{aligned}
$$

With this new information, player one should throw the deep pass 10 percent of the time and run up the gut 90 percent of the time. If a coach in the press box observed the defense as a 4-4 man, the updated perception about the action set of player two is

$$
\begin{aligned}
\hat{\beta}^{(2)} &= [\beta | \boldsymbol{\zeta}^{(2)} = \{\zeta_1 = QBObservation \cap \zeta_2 = CoachObservation\}] \\
&= [4 - 4Mannoblitz].
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\hat{\gamma}^{(2)} &= [\gamma | \hat{\beta}^{(2)}] \\
&= \{0, 0, 0, 1\}.
\end{aligned}
$$

3.7.4 Difference between Optimal and Perceived Optimal Decisions. The difference between these two updates can be used to measure the value of the next sensor observation, i.e., the coach observation. After the first update, the value of the

perceived optimal decision is

$$
\begin{aligned}
\hat{\pi}^{(1)} &= \hat{\gamma}^{(1)} \, \hat{\mathbf{R}}^{(1)} \, \hat{\delta}^{(1)'} \\
&= [0, 0.1, 0, 0.9]
\begin{bmatrix}
3.5 & 4.9 \\
4.2 & 15 \\
4 & 5.1 \\
4.4 & 3.2
\end{bmatrix}
\begin{bmatrix}
.9833 \\
.0167
\end{bmatrix} \\
&= 4.38
\end{aligned}
$$

while the perceived value of the second update is

$$
\begin{aligned}
\hat{\pi}^{(2)} &= \hat{\gamma}^{(2)} \, \hat{\mathbf{R}}^{(2)} \, \hat{\delta}^{(2)'} \\
&= [0, 0, 0, 1]
\begin{bmatrix}
3.5 \\
4.2 \\
4 \\
4.4
\end{bmatrix}
[1] \\
&= 4.4
\end{aligned}
$$

The perceived quality of the information gained can be measured by subtracting these two values. Thus the perceived value of ζ_2 is

$$
\hat{\pi}^{(2)} - \hat{\pi}^{(1)} = 4.4 - 4.38 = .02
$$

yards. The observation of the coach resulted in a perceived expected increase of .02 yards.

This may also be used in a post-game analysis using a tape of the game. Suppose the updated information strategy during the first update was $\hat{\gamma}^{(1)} = \{0, .1, 0, .9\}$ with the normal form given in Table 8, showing the quarterback perceiving the defense to be in a 4-4 or a 5-4. However, the true defensive formations that player two was selecting its strategy from was that in Table 9, a perceived optimal decision that

44

Table 8:   Updated Perceived 2-player Game

	4-4 Man no blitz	5-4 Man blitz
Option	3.5	4.9
Deep Pass	4.2	15
Short Pass	4	5.1
Run up the Gut	4.4	3.2

Table 9:   Updated True 2-player Game

	4-3 Prevent	4-3 Man
Option	9.4	4.75
Deep Pass	2.1	3.75
Short Pass	7.3	4.5
Run up the Gut	4.9	4.1

differs from the true optimal decision occurs. The true value of the perceived decision is

$$
\begin{aligned}
\hat{\pi}^{(1)} &= \hat{\gamma}^{(1)} \, \mathbf{R}^{(1)} \, \delta^{(1)\prime} \\
&= \begin{bmatrix} 0 & .1 & 0 & .9 \end{bmatrix} \begin{bmatrix} 3.5 & 4.9 \\ 4.2 & 15 \\ 4 & 5.1 \\ 4.4 & 3.2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= 4.0650
\end{aligned}
$$

while the true optimal decision would have resulted in a value of

$$
\begin{aligned}
\pi^{(1)} &= \gamma^{(1)} \, \mathbf{R}^{(1)} \, \delta^{(1)\prime} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 9.4 & 4.75 \\ 2.1 & 3.75 \\ 7.3 & 4.5 \\ 4.9 & 4.1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= 4.75
\end{aligned}
$$

Table 10:   Action Comparison

	4-3 Prevent	4-3 Man	4-4 Man no blitz	5-4 Man blitz	σ_i	$E(\alpha_i)$
Option	9.4	4.75	3.5	4.9	2.6	5.6
Deep Pass	2.1	3.75	4.2	15	5.9	6.3
Short Pass	7.3	4.5	4	5.1	1.5	5.2
Run up the Gut	4.9	4.1	4.4	3.2	0.7	4.2

The loss associated with the perceived optimal decision in this scenario is

$$\vec{\pi} = \hat{\pi}^{(1)} - \pi^{(1)}$$

$$= 4.0650 - 4.75$$

$$= -.685$$

yards. This value could be graphed over time to determine how well the quarterback is reading the defense or if the play calling is getting more accurate as the game is progressing.

3.7.5   Football Risk.   This section shows how the reward matrices change with differing levels of risk tolerance. Suppose the normal form of a football game is that in Table 10, where σ_i and $E(\alpha_i)$ are the standard deviation and expected outcome of the value of the game for action i computed across the action set of player two. The high risk action in the case of the football scenario is the deep pass evident by the high standard deviation. The lower risk actions for player one are the run up the gut and the short pass. Setting the risk tolerance of player one to that of an extremely risk averse individual, .5, results in the reward matrix

$$\mathbf{R} = \begin{bmatrix} 1 & .9950 & .9932 & .9963 \\ 0 & .9631 & .9850 & 1 \\ 1 & .9918 & .9776 & .9975 \\ .9963 & .9817 & .9899 & .8892 \end{bmatrix}$$

with strategy $\gamma = \{0, .0196, .8163, .1641\}$ and value $\pi=4.2425$. Setting the risk tolerance of player 1 to that of an extremely risk prone individual, -.5, results in the reward matrix

$$
\mathbf{R} = \begin{bmatrix}
0.00001367419 & 0.00000000124 & 0.00000000010 & 0.00000000168 \\
0.00000000000 & 0.00000000016 & 0.00000000041 & 1.00000000000 \\
0.00000020505 & 0.00000000075 & 0.00000000027 & 0.00000000251 \\
0.00000000168 & 0.00000000033 & 0.00000000061 & 0.00000000005
\end{bmatrix}
$$

with strategy $\gamma = \{.2498, .4955, .1355, .1192\}$ and value $\pi=4.2425$. The risk prone individual chooses to call the deep pass with a much higher probability.

3.8 Conclusion

Updating optimal decisions based on information available has received minimal attention, modeling this with game theory is an efficient technique to accomplish this. This chapter presented the methodology necessary to update decisions as more information becomes available, along with a method to measure the difference between perceived optimal decisions and true optimal decisions. The game theoretic techniques used to model a scenario were presented, and a procedure was given that accounts for the risk behavior of the players. Example calculations provided sufficient detail to demonstrate the application of these techniques. In chapter four, this methodology is applied to example scenarios and the consequences are examined.

IV. Results and Analysis

4.1 Introduction

The methodology presented in Chapter 3 is applicable in various fields of study. Combat situations, any naturally arising two player games such as those that occur in sports and recreation, strategic games, and proper allocation of resources are a few simple examples. Updating the optimal decision based on the reception of new information occurs in a broad range of areas; this chapter will demonstrate a few applications. Specifically, a combat scenario will be explored as well as a sports scenario. Finally, a resource allocation problem dealing with proper placement of funds to combat terrorist regimes is explored.

4.2 Utility

Each player will approach a situation in a different manner, in fact, their approach to a situation may vary as time progresses. For this reason, the utility of a value must be used to account for the decision maker preferences. This section sets forth the procedure for accomplishing this, as well as a technique to automate the risk behavior of a player based on his preferences in different types of situations. The concepts in the next section are from reference [11].

4.2.1 Certainty Equivalent. The certainty equivalent is a value that is used to determine ρ for the players of the game. The certainty equivalent may be found in several ways. The approach used in this research presents the decision maker with a proposition. The decision maker chooses the certain value that he prefers opposed to a gamble between two uncertain values. For example, the decision maker is faced with an uncertain gamble where 50% of the time he receives a value *low* and 50% of the time he receives a value *high*. The number he will he trade this gamble for is his certainty equivalent. If the number is the expected value of the gamble, $.5*low+.5*high = .5(low+high)$, the decision maker is labeled as risk neutral. If the number is less than the expected value, the decision maker is considered risk averse.

If the value is greater than the expected value, the decision maker is considered risk prone. The certainty equivalents are modified to obtain the standardized certainty equivalent $z_{.5}$ by using this equation:

$$z_{.5} = (CE - Low)/(High - Low). \tag{16}$$

$z_{.5}$ is then transformed into P by using Table 11. P is simply a standardized value of ρ. This value P is then multiplied by the range of the numbers considered, $range = low + high$. This value ρ is then plugged into the exponential utility function in Equation 10 on page 28 along with the value x that is being examined for utility. The utility of the number is then generated and used in place of the value in the original reward matrix $\mathbf{R}$.

4.2.2 Automating Rho. During an engagement where optimal decisions are being updated as information becomes available, it may be of interest to automate the risk attitude of a player towards a situation as it evolves. At the beginning of an engagement, a player may be risk averse but as the engagement progresses may decide to become more risk prone based on the actions of the other player or the situation of the game. To enumerate every type of situation that may occur during an engagement and ask the decision maker to determine a certainty equivalent for each one is unnecessarily exhaustive. A more efficient way to fully characterize the preference of the decision maker for each engagement is through the use of design of experiments. Initially, a screening experiment should be conducted to determine which factors play the biggest role in explaining the variability in the risk behavior of the players of the game. The high and low levels of the factors hypothesized to be important will be examined and a half factorial can be performed. After the most important factors are determined, the important levels of these factors can be examined through another designed experiment. Because of the fact that variance will not be present in the response of the decision maker (unless several decision makers are questioned), this resulting design can be used as a questionnaire to capture the

Table 11:   Certainty Equivalent Transformation

$z_{.5}$	P	$z_{.5}$	P	$z_{.5}$	P	$z_{.5}$	P
0	-	0.25	0.41	0.5	Inf	0.75	-0.41
0.01	0.01	0.26	0.44	0.51	-12.5	0.76	-0.39
0.02	0.03	0.27	0.46	0.52	-6.24	0.77	-0.36
0.03	0.04	0.28	0.49	0.53	-4.16	0.78	-0.34
0.04	0.06	0.29	0.52	0.54	-3.11	0.79	-0.32
0.05	0.07	0.3	0.56	0.55	-2.48	0.8	-0.3
0.06	0.09	0.31	0.59	0.56	-2.06	0.81	-0.29
0.07	0.1	0.32	0.63	0.57	-1.76	0.82	-0.27
0.08	0.12	0.33	0.68	0.58	-1.54	0.83	-0.25
0.09	0.13	0.34	0.73	0.59	-1.36	0.84	-0.24
0.1	0.14	0.35	0.78	0.6	-1.22	0.85	-0.22
0.11	0.16	0.36	0.85	0.61	-1.1	0.86	-0.2
0.12	0.17	0.37	0.92	0.62	-1	0.87	-0.19
0.13	0.19	0.38	1	0.63	-0.92	0.88	-0.17
0.14	0.2	0.39	1.1	0.64	-0.85	0.89	-0.16
0.15	0.22	0.4	1.22	0.65	-0.78	0.9	-0.14
0.16	0.24	0.41	1.36	0.66	-0.73	0.91	-0.13
0.17	0.25	0.42	1.54	0.67	-0.68	0.92	-0.12
0.18	0.27	0.43	1.76	0.68	-0.63	0.93	-0.1
0.19	0.29	0.44	2.06	0.69	-0.59	0.94	-0.09
0.2	0.3	0.45	2.48	0.7	-0.56	0.95	-0.07
0.21	0.32	0.46	3.11	0.71	-0.52	0.96	-0.06
0.22	0.34	0.47	4.16	0.72	-0.49	0.97	-0.04
0.23	0.36	0.48	6.24	0.73	-0.46	0.98	-0.03
0.24	0.39	0.49	12.5	0.74	-0.44	0.99	-0.01

decision makers risk preference at each combination of the factor levels. The design will not be used to explain variance in the system, but rather as an efficient way to automate ρ. Using the regression technique of least squares estimation,

$$\theta = (X'X)^{-1}X'y, \tag{17}$$

where X is the original design matrix and y is the response, a model of risk behavior is fit to the engagement. ρ_i, for $i = 1, 2$, can be automated based on any given situation during the engagement. Naturally, more data collected about the preferences of the decision maker will lead to a more accurate model of their actual behavior in each unique situation. For the purposes of this research, the levels of the important factors will be limited to a high and low case. In reality, many of the points between these high and low levels will be of interest and will assist in postulating an accurate model.

A study can also be performed on the risk behavior of the players to determine if they are approaching the situation in an optimal fashion. This can be done after the fact as a post analysis or before the situation occurs to develop an optimal strategy to approach the situation with.

4.2.2.1 Situation Vector. In these situations where the value of ρ can be automated, a situation vector describing the characteristics of the current situation and resulting value of ρ for player one is needed. Let,

$$S_t = \{X_1, X_2, \ldots, X_f, \rho_1, \rho_2\},$$

where t is the time step of the game, f is the number of important factors, X_i for $i = 1 : f$ is the level of the important factor i, ρ_1 is the risk tolerance of player one, and ρ_2 is the risk tolerance of player two. Notice t is different from s in Equation 4 on page 24 in that s is the updated time step for each observation and t is the overall time step of the game.

In situations where a game may only last a few plays and the risk tolerance of player one cannot be automated according to the situation, the risk tolerance notation will be accounted for on γ and δ. That is,

$$\hat{\gamma}_{\rho_1}^{(s)} = [\gamma | \{\hat{\beta}^{(s)} | \zeta^{(s)}\}]$$

and $\delta_{\rho_2}^{(s)}$.

4.2.3 Limitations of the Reward Matrix. When using the exponential utility function, the reward matrix is altered and extreme values may be used in calculating the strategies of the players to account for the risk behavior of the players. This can cause problems in the results if the reward matrix is not specified correctly. A short example should be suffice to demonstrate. Consider the following normal form of a game:

$P1 \downarrow \quad P2 \rightarrow$	Advance	Retreat
Advance	-3	4
Retreat	2	-2

When applying the exponential utility function to this matrix, the behavior of player one as the levels of the risk tolerance change are not entirely intuitive, although the reward matrix appears to be. See Table 12 for a display of the risk tolerance levels and the resulting strategies. Player one actually retreats more frequently as

Table 12:   Risk Tolerance Comparison of Original Matrix

Advance	0.1174	0.3288	0.3490	0.3612	0.3658	0.3721	0.3744	0.2655
Retreat	0.8826	0.6712	0.6510	0.6388	0.6342	0.6279	0.6256	0.7345
ρ	-1	-5	-10	-50	50	10	5	1

he becomes more risk prone, that is as his risk tolerance approaches 0 from negative infinity. Player one retreats more frequently as he becomes more risk averse as well,

that is as his risk tolerance approaches 0 from infinity. Even though these values are not intuitive, these are the actual values for this game setup. This is due to extreme values produced by the utility function. Consider the same game as above with normal form:

$P1 \downarrow \quad P2 \rightarrow$	Advance	Retreat
Advance	-4.2	2.1
Retreat	2	-.5

The values of the reward matrix have changed slightly from that in the original matrix, however these new values still make perfect sense. In fact, the orientation of the matrix has not changed, meaning $R_{1,1} < R_{1,2}$, $R_{2,1} > R_{2,2}$, etc. However compare the risk behavior using this new matrix by observing Table 13. As player one become more

Table 13:   Risk Tolerance Comparison of New Matrix

Advance	0.4542	0.3500	0.3191	0.2914	0.2767	0.2464	0.2076	0.0222
Retreat	0.5458	0.6500	0.6809	0.7086	0.7233	0.7536	0.7924	0.9778
ρ	-1	-5	-10	-50	50	10	5	1

risk prone, he advances with greater probability. As player one becomes more risk averse, he retreats most of the time. This is entirely intuitive, as opposed to the original matrix.

Now, suppose the orientation of the matrix actually changed but the matrix again still made sense in terms of the game. Instead of $R_{1,2} > R_{2,1}$, $R_{1,2} < R_{2,1}$. The risk strategies make even better sense with this updated matrix as shown in Table 14.

$P1 \downarrow \quad P2 \rightarrow$	Advance	Retreat
Advance	-4.2	1.8
Retreat	3	-2.5

These simple examples show the importance of correctly designating the reward matrix.

Table 14:   Risk Tolerance Comparison Changed Orientation Matrix

Advance	0.7682	0.5483	0.5139	0.4855	0.4710	0.4417	0.4046	0.1543
Retreat	0.2318	0.4517	0.4861	0.5145	0.5290	0.5583	0.5954	0.8457
ρ	-1	-5	-10	-50	50	10	5	1

4.2.3.1 Implications.   A key observation is that the expected case in all of these matrices is very similar. This shows that although two matrices can seem quite similar and produce similar results using traditional game theory, using the exponential utility function may accentuate the error in the matrices. This error is the difference between the constructed reward matrix and the true reward matrix. When the exponential transformations are applied, it is imperative that the constructed matrices be as accurate as possible. This is accomplished through using accurate data when it is available, or getting the adequate amount of surveys when constructing a rating type matrix. Small deviations in the reward matrix can cause major disruptions in the use of the exponential utility function. Yet, even though the strategies of player one are not intuitive in the above example, these are the actual strategies for this game setup. This is due to the extreme values produced by the utility function. The utility function actually accentuates any error present between the conjectured reward matrix and nature's truth. This error is actually in the relationships between the action sets of each player. See Tables 16 and 17 to see the values produced from the original game setup in Table 15 using the exponential utility function with risk tolerances of 1 and -1. Comparing Table 16 with Table 15, in the risk prone case with a risk tolerance of -1, the value of 4 to the decision maker is worth almost 10 times the value of 2. In the risk averse case, with a risk tolerance of 1, the value of 4 is worth essentially the same amount as the value of 2. An implication of this phenomenon is that a ranking type construction of the reward matrix may not be sufficient to produce accurate results when employing the exponential utility function. While using traditional game theory, the ranking system actually produces extremely intuitive results and is quite convenient; however it is invalidated for use

as the reward matrix when the exponential utility function is used to transform the matrix. A possible use of a ranking type system is to initially designate the orientation of the matrix. After ranking the possible outcomes from worst to best, actual values can be assigned for the reward matrix based on this initial orientation. This will eliminate the chance of composing a reward matrix that produces counterintuitive results.

Table 15:   Original Reward Matrix

RT = ∞	Advance	Retreat
Advance	-3	4
Retreat	2	-2

Table 16:   Risk Prone Transformed Reward Matrix

RT = -1	Advance	Retreat
Advance	0	1
Retreat	.1345	.0016

Table 17:   Risk Averse Transformed Reward Matrix

RT = 1	Advance	Retreat
Advance	0	1
Retreat	.9942	.6327

4.3   Combat Scenario

Modeling combat situations between an entity and nature using game theory is an effective modeling technique for use in simulation models or during and after an actual battle. The quality of the information input into the game will have a direct effect on the quality of the outputs of the simulation. Thus, the inputs must be representative of the true combat scenarios in order to produce accurate combat situations or games.

This section first presents the initial setup of an example combat game. Next, an analysis of the number of surveys collected **N** and how this directly affects the outputs

of the combat game is studied. This analysis will provide the adequate number of surveys to collect that will verify the use of this methodology in a combat scenario. The methodology will then be applied to an example combat game and a proper analysis conducted.

4.3.1 Game Setup. Initially, the action sets $\boldsymbol{\alpha}$ of player one and $\boldsymbol{\beta}$ of nature must be brainstormed by knowledgeable decision makers. Suppose that after thinking about a situation in which a tank observes an object in the distance,

$$\boldsymbol{\alpha} = [ShootMortar, Advance, DoNothing, Communicate]$$

while

$$\boldsymbol{\beta} = [EnemyTruck, CivilianTruck, EnemyTank,$$
$$EnemyArmoredPersonnelCarrier, FriendlyTank]$$

in the one player versus nature case, where nature, $\boldsymbol{\beta}$, is the sensor inputs to player one. Recall Table 1 on page 19 where the combinations of the action sets are assigned a rank based on a scale of severity from -5 to 5. A survey is given to expert operators questioning what the outcome is for the situations in Table 18. The table shows a hypothetical response by one subject matter expert. Keep in mind throughout this example that in the one player versus nature game, we only need to be concerned with the strategy of player one.

4.3.1.1 Assumptions. Some assumptions must be made and presented to the SME's in order to ensure stability of the responses. For example, one SME could assume the civilian truck may possibly be a suicide bomber or an innocent civilian while another assumes it just to be an innocent civilian. This would result in extreme differing values in the reward matrix. These assumptions should be clearly stated whenever a reward matrix is being formulated.

Table 18:   User Survey Data

Player 1	Nature	Reward
Shoot Mortar	Enemy truck	3
Advance	Enemy truck	-3
Do nothing	Enemy truck	-2
Communicate	Enemy truck	2
Shoot Mortar	Civilian Truck	-4
Advance	Civilian Truck	5
Do nothing	Civilian Truck	3
Communicate	Civilian Truck	-3
Shoot Mortar	Enemy Tank	4
Advance	Enemy Tank	-5
Do nothing	Enemy Tank	-2
Communicate	Enemy Tank	3
Shoot Mortar	Enemy APC	5
Advance	Enemy APC	-4
Do nothing	Enemy APC	-3
Communicate	Enemy APC	4
Shoot Mortar	Friendly Tank	-5
Advance	Friendly Tank	5
Do nothing	Friendly Tank	3
Communicate	Friendly Tank	-3

- Normal battle scenario based in the desert

- Tank observes an unknown object with onboard sensors

- Tank is at war with a hostile enemy

- Mission is to destroy enemies on sight

- Trying to minimize civilian casualties

- Enemy truck houses men with weapons

- Civilian truck is innocent

- Communication implies radioing for backup

Stating the assumptions up front ensures the SME's fully understand the scenario.

4.3.1.2 Survey Response Effects.   Each SME will differ slightly in their opinion of the outcomes in Table 18. As the number of surveys **N** approaches infinity,

the reward matrix $\mathbf{R}$ will approach the true reward matrix. The number of surveys collected will have an impact on the quality of the information gleaned from this combat game. The best comparison parameter is the value of the game π to player 1. In Figure 2, the simulated value of the game approaches the true value of the game, $\pi = .0588$, as the number of survey responses increases. The simulation was generated using the Matlab programming language. Intuitive response variation was assigned to the responses of the SME's.
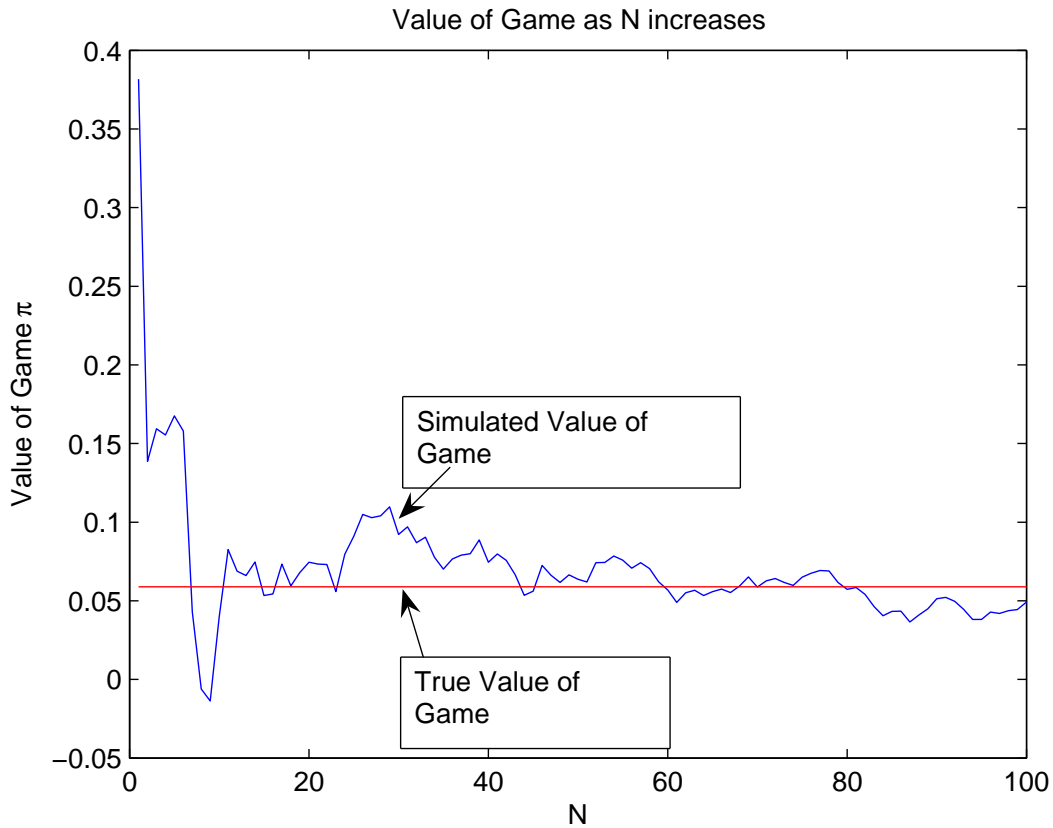

Figure 2:   Effect of $\mathbf{N}$ on Response

4.3.2   Running the Game.   After receiving and averaging the survey responses from the subject matter experts, the normal form of the game is given in Table 19. Keep in mind, the values here are just rated values so their meanings are only relative.

Table 19:   Normal Form of Combat Game

	Enemy truck	Civilian Truck	Enemy Tank	Enemy APC	Friendly Tank
Shoot Mortar	2.6	-4.6	4.7	5	-4.1
Advance	-2.7	4.9	-5	-3.7	5
Do nothing	-1.2	1.6	-1.9	-3.2	3.6
Communicate	1.7	-1.9	2.8	2.5	-1.6

Initially, the game must be examined for a saddle point, this particular game does not have a saddle point and must be solved using linear programming.

$$\max z = v + 0w_1 + 0w_2 + 0w_3 + 0w_4$$

$$s.t. \; v \leq 2.6w_1 - 2.7w_2 - 1.2w_3 + 1.7w_4$$

$$v \leq -4.6w_1 + 4.9w_2 + 1.6w_3 - 1.9w_4$$

$$v \leq 4.7w_1 - 5w_2 - 1.9w_3 + 2.8w_4$$

$$v \leq 5w_1 - 3.7w_2 - 3.2w_3 + 2.5w_4$$

$$v \leq -4.1w_1 + 5w_2 + 3.6w_3 - 1.6w_4$$

$$\sum_i w_i = 1$$

$$w_i \geq 0 \; \forall \, i.$$

which yields the mixed strategy

$$\hat{\gamma}^{(0)} = \{w_1, w_2, w_3, w_4\}$$
$$= \{0, .3214, 0, .6786\}.$$

This is the expected case, the probability distribution player one should follow in this game if he is risk neutral. 32% of the time he should advance further and 67% of the time he should communicate. This may not always be the strategy that player one chooses to run because of the manner in which he views the game. Table 20 shows a

comparison of the strategies associated with different risk behaviors. Recall the risk behavior associated with the values of ρ:

-1 - Extremely Risk Prone

-10 - Moderately Risk Prone

$|\infty|$ - Expected Case or Risk Neutral

10 - Moderately Risk Averse

1 - Extremely Risk Averse

If a player is an extremely risk prone individual, his mortar is shot 91 percent of the time and he advances with probability .09. If the player is extremely risk averse, 79 percent of the time he should communicate for backup and 21 percent of the time he should do nothing, but he should never shoot or advance.

Table 20:   Risk Behavior Comparison

Rho	-1	-10	Infinity	10	1
Player 1 Actions					
Shoot Mortar	0.9089	0.5662	0	0	0
Advance	0.0911	0.4338	0.3214	0.2123	0
Do nothing	0	0	0	0.1694	0.2135
Communicate	0	0	0.6786	0.6184	0.7865
Value of Game (π)	-0.1817	-0.0055	0.2857	0.2204	0.2034

The effects of the risk behavior of player one can be studied in relation to the strategy of player two. As shown in Table 20, the optimal strategy, noted by the value of the game, is always the expected case, player one is maximizing his minimum gain. This is true whenever it is hypothesized that player two is using the maximin principle. Any deviation from this behavior causes a decrease in the expected value of the game to player one. In this case, player two is nature and thus will not always use the maximin principle, there are number of different strategies that may occur. Therefore it doesn't make sense to compare the value of the game for different risk strategies of the two players of the game. This is only useful during the two-player

60

game. The main study then is the variation in the response due to the selection of a certain risk behavior given a distribution on the probable outcomes of nature. This will be accomplished during the post-war analysis.

Next, suppose the tank receives information from its sensors that the sighted object is a tank of some sort, a tracked vehicle. The action set player one perceives nature to be choosing from is dependent on his sensors. Recall Equation 5 on page 24,

$$
\begin{aligned}
\hat{\beta}^{(1)} &= [\beta | \zeta^{(1)} = \{\zeta_1 = OnboardSensor\}] \\
&= \{EnemyTank, EnemyAPC, FriendlyTank\}
\end{aligned}
$$

The strategy of player one will update due to his perception of the action set of player two, that is

$$
\begin{aligned}
\hat{\gamma}^{(1)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{0, .2437, .0899, .6663\}.
\end{aligned}
$$

The percentage of time player one advances is decreased due to his gained knowledge, and the percentage of time he does nothing and waits is increased. This makes sense because if he knows the object is most likely a tank of enemy nature (uniform distribution raised from .2 to .33), he may incur more damage by advancing. However, the risk prone individual approaches the situation by either advancing or shooting his mortar,

$$
\hat{\gamma}^{(1)}_{(-1)} = \{.5745, .4255, 0, 0\}.
$$

The risk prone individual will take extreme measures to accomplish his mission. Next, suppose the tank received information from an airborne reconnaissance source that the tracked vehicle was heavily armored, indicating that the object was indeed a tank

but not an armored personnel carrier. Now,

$$
\begin{aligned}
\hat{\beta}^{(2)} &= [\beta | \zeta^{(2)} = \{\zeta_1 = OnboardSensor \cap \zeta_2 = AirborneReconnaissance\}] \\
&= [EnemyTank, FriendlyTank]
\end{aligned}
$$

with a resulting strategy

$$
\hat{\gamma}_{(0)}^{(2)} = \{0, 0, .4444, .5556\}.
$$

Since player one is unsure the identity of the tank, it makes sense for him to do nothing and wait for more information or communicate for backup from friendly forces.

Finally, suppose the tank receives visual confirmation from a special forces troop that the object is indeed an enemy tank. The perceived optimal strategy is based on

$$
\begin{aligned}
\hat{\beta}^{(3)} &= [\beta | \zeta^{(3)} = \{\zeta_1 = OnboardSensor \cap \zeta_2 = AirborneReconnaissance \cap \zeta_3 = SpecialForces\}] \\
&= [EnemyTank]
\end{aligned}
$$

and results in the strategy

$$
\hat{\gamma}_{(0)}^{(3)} = \{1, 0, 0, 0\}.
$$

Player one will always shoot his mortar in this situation based on his perception that the object is an enemy tank.

Consider now the case where ζ_3, the special forces troop, was in error regarding the identity of the object. In fact, the information received up to the point of the special forces input was correct, however the special forces troop mistakenly failed to identify the object as a friendly tank. Thus the true optimal strategy of the game is

$$
\gamma_{(0)}^{(3)} = \{0, 1, 0, 0\},
$$

player one should advance every time in the situation where the object is a friendly tank. The regret of the perceived optimal strategy can be measured using Equa-

tion 9 on page 26, where

$$\hat{\pi}^{(3)} = \hat{\gamma}^{(3)} \mathbf{R}^{(3)} \delta^{(3)'}$$

$$= [1,0,0,0] \begin{bmatrix} -4.1 \\ 5 \\ 3.6 \\ -1.6 \end{bmatrix} [1]$$

$$= -4.1$$

is the perceived optimal value of the game given the truth and

$$\pi^{(3)} = \gamma^{(3)} \mathbf{R}^{(3)} \delta^{(3)'}$$

$$= [0,1,0,0] \begin{bmatrix} -4.1 \\ 5 \\ 3.6 \\ -1.6 \end{bmatrix} [1]$$

$$= 5$$

is the true optimal value of the game. Equation 9 on page 26 yields the following regret for player one because of his decision:

$$\vec{\pi}^{(3)} = \hat{\pi}^{(3)} - \pi^{(3)}$$

$$= (-4.1) - 5$$

$$= -9.1.$$

This is a very high regret as would be the case if a mortar was shot at a friendly tank.

Player two (nature) will always be only one of the objects in the original action set. This is why the perceived optimal strategy is multiplied by the column in the reward matrix corresponding to the true identity of the object. This gives the true value of the perceived optimal strategy.

4.3.3 Post-Battle Analysis. The methodology in this research can be applied in hindsight to provide feedback on the performance of the sensors of a tank and intuition on future strategies to approach similar situations with.

The initial game can be examined for insight on the possible strategies that player one should take in the future. In Figure 3, it is observed that by player one choosing to approach the initial game in an extreme risk prone manner, he can expect to gain more value than by using the expected case. This is based on the assumption that the probabilities of the outcome of nature is uniform, $\delta\,(U)$ for β.

The best way to explore the consequences of different risk behavior on the outcome of the game is through examining all of the possible outcomes and noting the value that each strategy produces at each of the possible outcomes. This can be accomplished through exploration of the response surface, observing the expected gain at each of the risk strategies, or looking individually at each interaction plot. Table 21 shows the risk strategies of player one and the possible outcomes of the situation, or the moves of nature for the initial game. These values are calculated by multiplying the strategy produced by the risk tolerance of player one with the sure outcome of nature column of R. For example, with $\rho = -1$, the strategy of player one is

$$\gamma_{(-1)} = \{.9089, .0911, 0, 0\}.$$

The first column of R is used to calculate the value of this strategy when the truth of player two is the Enemy Truck. The value of the risk prone strategy when player

Table 21:   Risk Tolerance Comparison

Risk Tolerance (ρ)	-1	-2	-5	-10	-20	-50
Enemy Truck	2.1172	1.3303	0.5786	0.3011	0.3470	0.3094
Civ Truck	-3.7345	-2.3242	-0.9767	-0.4792	0.1909	0.2492
Enemy Tank	3.8163	2.3762	1.0004	0.4925	0.4016	0.3348
Enemy APC	4.2074	2.9158	1.6818	1.2263	0.5936	0.5405
Friendly Tank	-3.2710	-1.9200	-0.6292	-0.1528	0.4294	0.4860
Mean	0.6271	0.4756	0.3310	0.2776	0.3925	0.3839
StDev	3.8545	2.4431	1.1142	0.6528	0.1454	0.1235
Risk Tolerance (ρ)	50	20	10	5	2	1
Enemy Truck	0.2944	0.2490	0.2749	0.3505	0.7199	1.0809
Civ Truck	0.2724	0.1996	0.1362	-0.0107	-0.7171	-1.1528
Enemy Tank	0.3082	0.2920	0.3483	0.4959	1.2115	1.7967
Enemy APC	0.5193	0.2207	0.2186	0.2759	0.5735	1.2832
Friendly Tank	0.5085	0.7001	0.6816	0.5942	0.1575	-0.4899
Mean	0.3805	0.3323	0.3319	0.3411	0.3891	0.5036
StDev	0.1225	0.2085	0.2103	0.2324	0.7241	1.2594

two is an Enemy Truck is

$$[.9089, .0911, 0, 0] \begin{bmatrix} 2.6 \\ -2.7 \\ -1.2 \\ 1.7 \end{bmatrix} [1] = 2.1172.$$

The mean of the risk prone approach is calculated assuming a uniform distribution of the actions of nature. This is a valid assumption with prior information unavailable. The mean and standard deviation are calculated across all the possible actions of nature for each action of player one. This gives the expected reward player one can gain along with the amount of variation expected for each risk strategy of player one.

Table 21 shows that as the risk tolerance of player one approaches the risk neutral or expected case, the expected value of the game decreases as well as the standard deviation. As player one becomes more risk neutral, he can expect to achieve low variation in the value of the game, but as he becomes more risk prone, he can expect a much larger variation in the value of the game. Figure 3 shows the mean and variance of the initial scenario. The variance increases as player one becomes more risk prone or risk averse. This type of chart can be used to weigh tradeoffs of using different risk strategies. For instance, if player one was constrained to gaining a certain amount of value but the chance for loss needed to minimal, the expected case may be the best choice. However, if the player could afford a possible loss for a greater gain, he may choose the risk prone approach. This type of analysis can be performed for each
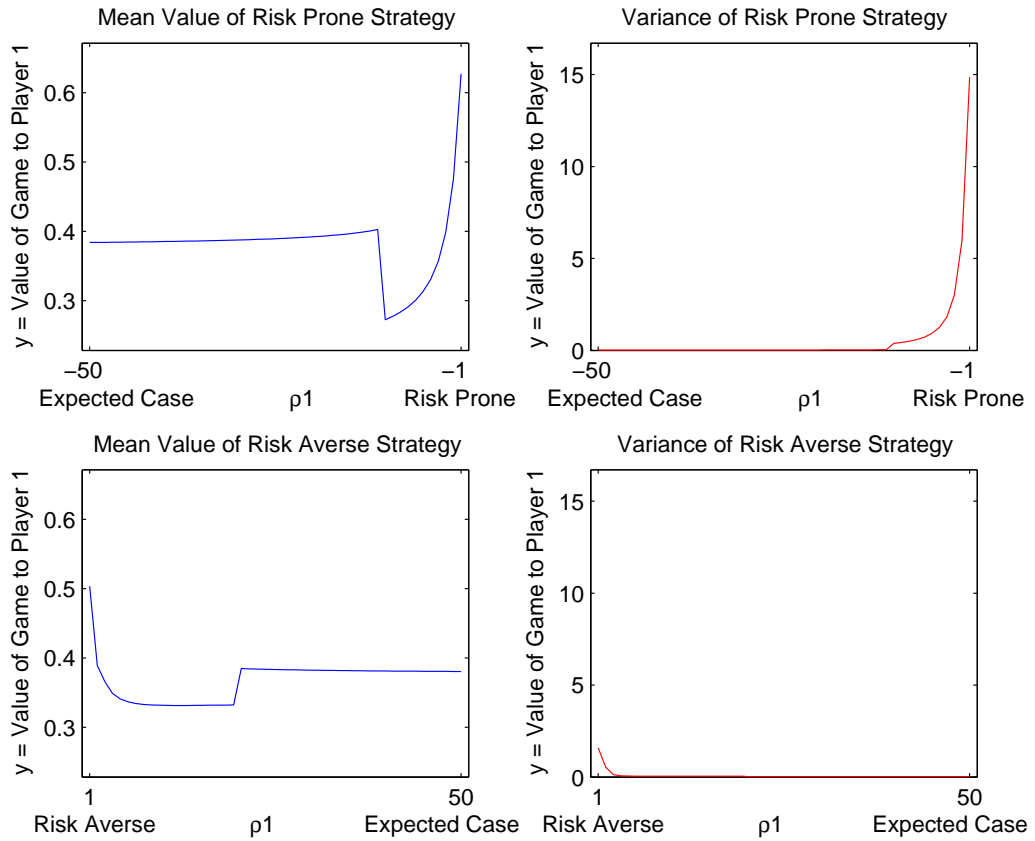

Figure 3: Effects of ρ on Game Value

update of the game, and will be different for each game.

Suppose now that *a priori* information was available regarding the possible outcomes of nature, the best strategy for player one will change. Consider the same example above with an *a priori* distribution on the actions of nature such that the probabilities of the object are:

$$
\left\{
\begin{array}{ll}
EnemyTruck, & .2; \\
CivilianTruck, & .05; \\
EnemyTank, & .5; \\
EnemyAPC, & .2; \\
FriendlyTank, & .05.
\end{array}
\right\}.
$$

Figure 4 shows the updated mean and variance plot. Table 22 shows the updated comparisons as well. The mean increased significantly in the risk prone case with the updated probabilities, while the variance stayed roughly the same. It makes sense that when there is a higher probability that the object is an enemy, risk prone behavior will be more beneficial. There is still a chance of an extremely bad outcome though shown by the large variance.

Consider one more case where the information known about the distribution of the actions of nature give the probabilities:

$$
\left\{
\begin{array}{ll}
EnemyTruck, & .1; \\
CivilianTruck, & .4; \\
EnemyTank, & .05; \\
EnemyAPC, & .1; \\
FriendlyTank, & .35.
\end{array}
\right\}.
$$

Figure 5 and Table 23 show the updated mean and variance plots and comparison values. The expected gain for player one is very low if he chooses to play risk prone with prior probabilities indicating the object is most likely a friendly of
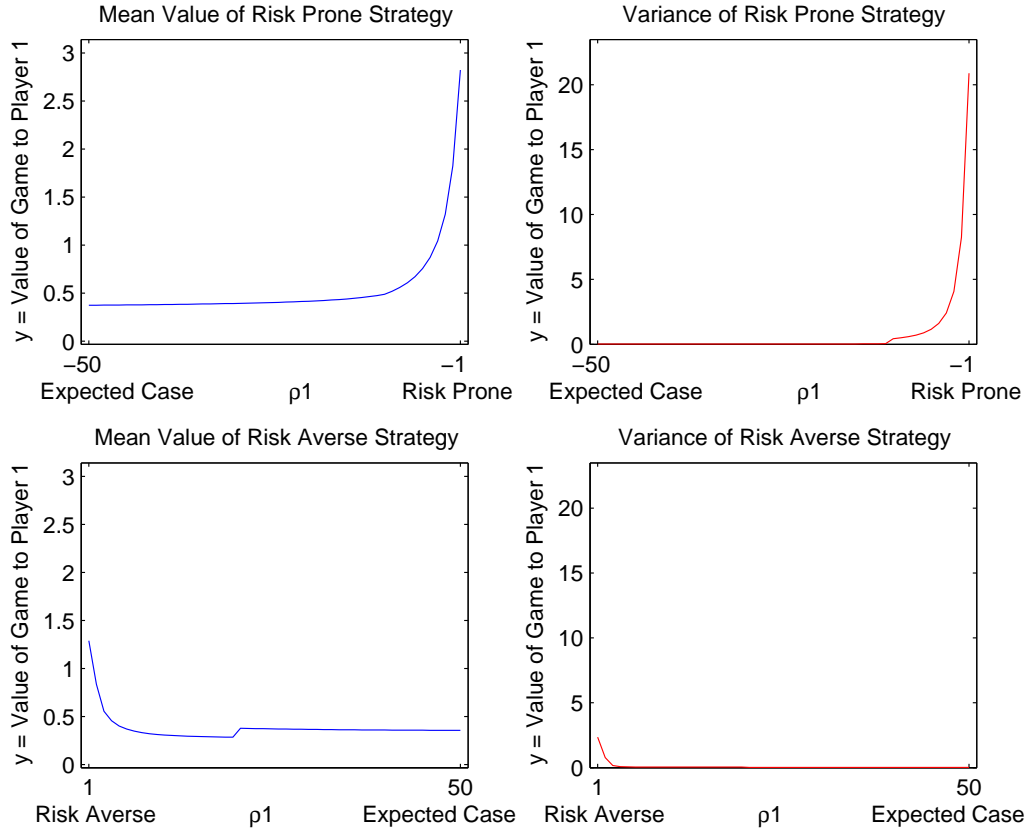
Figure 4:   Effects of ρ on Game Value

some sort. The variance stays roughly the same indicating that it is still possible to gain a great deal by being risk prone, just more unlikely than in the above cases. In this scenario, the best risk strategy is the risk neutral expected case. This ensures a certain expected reward with virtually no variation.

Again, this can be explored for each update of the reward matrix during a game. The scenarios will differ with different games and with different a priori distributions. There are numerous possibilities here, only the surface has been scratched with the above examples.

Table 22:   Updated Risk Tolerance Comparison

Risk Tolerance	-1	-2	-5	-10	-20	-50
Enemy Truck	2.1172	1.3303	0.5786	0.3011	0.3470	0.3094
Civ Truck	-3.7345	-2.3242	-0.9767	-0.4792	0.1909	0.2492
Enemy Tank	3.8163	2.3762	1.0004	0.4925	0.4016	0.3348
Enemy APC	4.2074	2.9158	1.6818	1.2263	0.5936	0.5405
Friendly Tank	-3.2710	-1.9200	-0.6292	-0.1528	0.4294	0.4860
Mean	2.8228	1.8251	0.8720	0.5201	0.4199	0.3741
StDev	4.5699	2.8714	1.2678	0.7069	0.1486	0.1240

Risk Tolerance	50	20	10	5	2	1
Enemy Truck	0.2944	0.2490	0.2749	0.3505	0.7199	1.0809
Civ Truck	0.2724	0.1996	0.1362	-0.0107	-0.7171	-1.1528
Enemy Tank	0.3082	0.2920	0.3483	0.4959	1.2115	1.7967
Enemy APC	0.5193	0.2207	0.2186	0.2759	0.5735	1.2832
Friendly Tank	0.5085	0.7001	0.6816	0.5942	0.1575	-0.4899
Mean	0.3559	0.2849	0.3137	0.4024	0.8365	1.2890
stdev	0.1255	0.2151	0.2113	0.2423	0.8801	1.5353
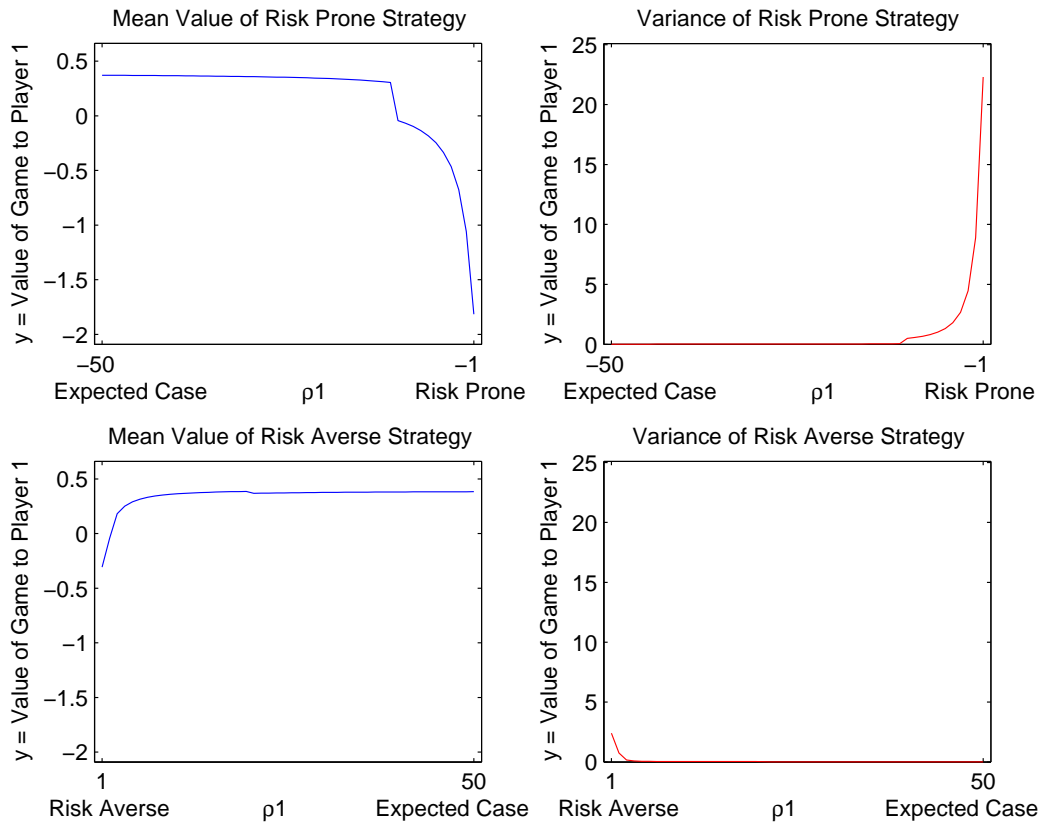
Figure 5:   Effects of ρ on Game Value

Table 23:   2nd Updated Risk Tolerance Comparison

Risk Tolerance	-1	-2	-5	-10	-20	-50
Enemy Truck	2.1172	1.3303	0.5786	0.3011	0.3470	0.3094
Civ Truck	-3.7345	-2.3242	-0.9767	-0.4792	0.1909	0.2492
Enemy Tank	3.8163	2.3762	1.0004	0.4925	0.4016	0.3348
Enemy APC	4.2074	2.9158	1.6818	1.2263	0.5936	0.5405
Friendly Tank	-3.2710	-1.9200	-0.6292	-0.1528	0.4294	0.4860
Mean	-1.8154	-1.0582	-0.3348	-0.0678	0.3408	0.3715
StDev	4.7238	2.9849	1.3400	0.7585	0.1565	0.1243

Risk Tolerance	50	20	10	5	2	1
Enemy Truck	0.2944	0.2490	0.2749	0.3505	0.7199	1.0809
Civ Truck	0.2724	0.1996	0.1362	-0.0107	-0.7171	-1.1528
Enemy Tank	0.3082	0.2920	0.3483	0.4959	1.2115	1.7967
Enemy APC	0.5193	0.2207	0.2186	0.2759	0.5735	1.2832
Friendly Tank	0.5085	0.7001	0.6816	0.5942	0.1575	-0.4899
Mean	0.3837	0.3864	0.3598	0.2911	-0.0418	-0.3064
stdev	0.1225	0.2171	0.2126	0.2391	0.8697	1.5511

4.4   *Sports Application*

This methodology fits nicely to a football game where the offense is attempting to audible plays based on the observations of the quarterback or coaches. Initially, the game can be set up for each situation of the game. As the quarterback approaches the line, and the coach observes the defense from the sidelines or press box, the defensive formation can be estimated thus eliminating some of the possible defensive setups. This leads to an updated offensive strategy that is based on this perception. The risk behavior of the teams can also be estimated and will change with each play of the game.

4.4.1   Initial Game Setup.   Initially, the plays that are available to each team must be determined for various situations during the game. For instance, when the offense is within 10 yards of the opponents endzone, the long pass is not a possible action to call. Most plays will be available the majority of the game. After the action sets are determined for the offense and defense, the proper statistics must be gathered from past games. Each situation where the offensive action has been used against the defensive formation must be assigned an average number of yards gained from statistical data. Table 24 shows an example of the normal form of a football game after data collection. The defensive formations are designed to limit certain plays, here are the plays that are best defended against by each formation:

- 4-4 Overload - Sweep

- 5-4 Blitz - Middle Run

- 4-4 Zone - Short Pass

- 4-3 Man - Long Pass.

4.4.2   Automating Rho.   In determining the general risk behavior of a coach, the initial factors that cause a coach to vary his play calling according to the amount of risk he is willing to accept must be expounded. After brainstorming all the possible

Table 24:   Initial Football Game

	4-4 Overload	5-4 Blitz	4-4 Zone	4-3 Man
Sweep	-2.7	3.8	4	6.6
Middle Run	3.4	2	5.1	5.9
Short Pass	4	3.9	0.5	7.3
Long Pass	6.1	7	6.3	-3.4

factors that could affect risk behavior during a football game, design of experiments can be used to determine the most influential factors, the factors that cause the variation in the response variable certainty equivalent. A two-level fractional factorial can be used to weed out the unimportant factors. Suppose that this process revealed the most important factors as down, distance to go for a first down, field position, time left in the game, and score of the game. For simplicity, the scenarios presented herein assume the score and a time left in the game. Let's also assume for the sake of brevity that the remaining factors, down, distance, and field position, only possess two levels, high and low. In reality, there may be four or more levels for each of the original factors, the more the better. Table 25 shows a description of the high and low levels of the three factors.

Table 25:   Factor Levels

Factor	High(+)	Low(-)
Field Position	≥ 50	< 50
Down	$\geq$ 3rd	$\leq$ 2nd
Distance to Flag	≥ 8	< 8

Table 26 shows a simple setup of a design matrix that allows ρ to be automated according to down, distance, and field position by inputting a certainty equivalent for each design point. The decision maker is asked to answer a question at each of the design points in Table 26. The question is the certain number of yards willing to be accepted as a trade for the gamble: 50% chance of gaining 3 yards and 50% chance of gaining 10 yards. The responses are given, if the decision maker chooses the expected value, 6.5 in this case, he is considered a risk neutral individual. A value of larger than 6.5 implies risk prone and less than 6.5 indicates a risk averse attitude. The

73

values of the certainty equivalent have been formulated under the assumption that player one is losing by 7 points with 5 minutes or less remaining in the game.

Table 26:   User Risk Survey

Field Position(X_1)	Down(X_2)	Distance(X_3)	Certainty Equivalent(y)
+	+	+	8
+	+	-	6.3
+	-	+	6
+	-	-	5
-	+	+	9.5
-	+	-	8.5
-	-	+	8
-	-	-	7.5

While fitting a model to the input data, it is important to use the certainty equivalent as the response. If the actual value of ρ is used, the model may need cubic terms or higher, which will make the automation process much more difficult and time consuming. After the model has been fit using the certainty equivalent, the values of ρ can be calculated. An accurate model is fit using just main effects and interaction terms. The design matrix $\mathbf{X}$, initial y response vector, fitted values $\hat{y}$, standardized CE $z_{0.5}$, standardized ρ value P, and corresponding ρ are given in Table 27. $z_{.5}$ is computed using the certainty equivalent in Table 26 and Equation 16. For example, $z_{.5}$ for the first combination of factors in Table 26 is

$$
\begin{aligned}
z_{.5} &= (CE - Low)/(High - Low) \\
&= (8 - 3)/(10 - 3) \\
&= .714
\end{aligned}
$$

where high and low are the respective values of the lottery given to the decision maker. Note, ρ is calculated by plugging $z_{.5}$ into Table 11, obtaining a standardized value of ρ, P, and then multiplying this by the $range = high - low$. For the first design point, $z_{.5} = .714$ corresponds to an P value of -.52. This is multiplied by the range, $range = high - low = 7$, to return $\rho = -3.64$. The values for $\hat{y}$ were obtained using

74

Table 27:   Design Matrix for Automating ρ

Intercept	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y	$\hat{y}$	$z_{0.5}$	P	ρ
1	1	1	1	1	1	1	8	7.975	0.714285714	-0.52	-3.64
1	1	1	-1	1	-1	-1	6.3	6.325	0.471428571	4.16	29.12
1	1	-1	1	-1	1	-1	6	6.025	0.428571429	1.76	12.32
1	1	-1	-1	-1	-1	1	5	4.975	0.285714286	0.52	3.64
1	-1	1	1	-1	-1	1	9.5	9.525	0.928571429	-0.1	-0.7
1	-1	1	-1	-1	1	-1	8.5	8.475	0.785714286	-0.32	-2.24
1	-1	-1	1	1	-1	-1	8	7.975	0.714285714	-0.52	-3.64
1	-1	-1	-1	1	1	1	7.5	7.525	0.642857143	-0.85	-5.95

the least squares technique from Equation 17 which yielded the following model:

$$
\begin{aligned}
\hat{y} &= \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3 + \theta_{12} X_1 X_2 + \theta_{13} X_1 X_3 + \theta_{23} X_2 X_3 \\
&= 7.35 - 1.025 X_1 + .725 X_2 + .525 X_3 + .1 X_1 X_2 + .15 X_1 X_3 + .15 X_2 X_3.
\end{aligned}
$$

This may not seem useful at first glance, however when many variables are present and several levels exist for each variable, it is imperative to have a prediction equation. In this case the levels of the variables are categorical, either high or low. When the levels of the factors are continuous, the time of game and the yardline for instance, it is important to have the ability to predict between design points. Again, the above equation is just the certainty equivalent for the situation where player one is losing by 7 points with 5 minutes or less remaining. Ideally, all of the factors will be included in the model and the risk tolerance for every conceivable situation spanning the entire length of the game will be approximated and automated.

Risk behavior can now be automated for any field position, down, and distance to the first down according to the risk preferences of the decision maker, the coach in this football scenario.

4.4.3 Running the Game.   With the data from the initial game setup and the risk preference function of the decision maker, the game commences. With 5 minutes remaining in the game and losing by 7 points, the offense is facing 3rd and 10

on the -38 yardline (a negative sign in front of the yardline implies the offenses own half of the field, whereas no sign implies the defensive half of the field). There are 60 minutes in an entire football game, the number of minutes are thus the time that remains in the game. ρ_1 is the risk tolerance of player one according to the inputs of the situation of the game. ρ_2 is the estimation of the risk tolerance of player two by player one. We will assume for this scenario that player two (defense) knows the risk tolerance of player one (offense) in most situations that arise during a football game. The case where his assumption is wrong will be addressed further in upcoming sections. All of these inputs combined are considered the situation of the game, **S**.

$$
\begin{aligned}
\mathbf{S_1} &= \{Score, TimeRemaining, FieldPosition, Down, Distance, \rho_1, \rho_2\} \quad (18) \\
&= \{-7, 5, -38, 3, 10, -.7, -.7\}
\end{aligned}
$$

shows the situation during the first play of the game. Obviously, the game will normally begin with a score of 0-0 and 0 time elapsed. The game is started here to show the application of the automation of ρ.

The risk tolerance of player one in (18) was found using the risk tolerance function and Table 25. The offensive certainty equivalent in this situation is

$$
\begin{aligned}
\hat{y} &= 7.35 - 1.025X_1 + .725X_2 + .525X_3 + .1X_1X_2 + .15X_1X_3 + .15X_2X_3 \\
&= 7.35 - 1.025(-1) + .725(1) + .525(1) + .1(-1) + .15(-1) + .15(1) \\
&= 9.525.
\end{aligned}
$$

This number is converted to ρ using the procedure outlined above. Player one has a risk tolerance of $\rho = -.7$, which is an extreme risk prone approach to the situation. Initially, in the huddle, with lack of prior information about the defensive formation $\boldsymbol{\beta}$ that player two will call, player one chooses from his action set

$$
\boldsymbol{\alpha} = [Sweep, MiddleRun, ShortPass, LongPass]
$$

based on his perception of the available plays that player two can choose from.

$$\hat{\gamma}_{\mathbf{S_1}}^{(0)} = [\gamma = \{0, 0, .1538, .8462\} \mid$$

$$\hat{\beta}^{(0)} = \{4 - 4Overload, 5 - 4Blitz, 4 - 4Zone, 4 - 3Man\},$$

the initial strategy of player one based on his risk tolerance during the situation and his perception of all the available plays to player two. This shows that because of the risk prone behavior of player one due to the situation, he will call the long pass 85% of the time and the short pass 15% of the time. The strategy of player two is

$$\begin{aligned} \hat{\delta}_{\mathbf{S_1}}^{(0)} &= \{\delta | \hat{\beta}_{\mathbf{S_1}}^{(0)}\} \\ &= \{.8539, 0, 0, .1461\}. \end{aligned} \tag{19}$$

Player one believes player two will choose to run the 4-4 Overload or the 4-3 Man. This makes sense looking at the normal form of the original game in Table 24 because a risk prone attitude by player two will cause him to want to gain and not worry about losing. The two negative numbers in the original reward matrix correspond to those two actions.

As the quarterback approaches the line of scrimmage, he observes the defense in either a 4-3 man or a 4-4 zone. The perceived actions available to player two are dependent on the quarterback perception:

$$\begin{aligned} \hat{\beta}_{\mathbf{S_1}}^{(1)} &= [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = QBObservation\}] \\ &= [4 - 4Zone, 4 - 3Man]. \end{aligned} \tag{20}$$

His strategy thus updates based on this observation,

$$\begin{aligned} \hat{\gamma}_{\mathbf{S_1}}^{(1)} &= \{\gamma | \hat{\beta}_{\mathbf{S_1}}^{(1)}\} \\ &= \{0, 0, .1933, .8067\}. \end{aligned} \tag{21}$$

Even though player one knows that player two is defending heavily against the pass by playing the 4-3 Man and the 4-4 Zone, he will still call a pass because he is in a situation where he must get yards and a first down or he will lose the game. Player one now chooses a play according to this distribution and possibly calls an audible to his original play out of the huddle. Based on his knowledge of the actions available to player two and the strategy of player two

$$\hat{\delta}_{\mathbf{S_1}}^{(1)} = \{.8067, .1933\}, \tag{22}$$

he expects to gain

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_1}}^{(1)} &= \hat{\gamma}_{\mathbf{S_1}}^{(1)} \, \hat{\mathbf{R}}_{\mathbf{S_1}}^{(1)} \, \hat{\delta}_{\mathbf{S_1}}^{(1)\prime} \\
&= [0, 0, .1933, .8067]
\begin{bmatrix}
4 & 6.6 \\
5.1 & 5.9 \\
0.5 & 7.3 \\
6.3 & -3.4
\end{bmatrix}
\begin{bmatrix}
.8067 \\
.1933
\end{bmatrix} \\
&= 3.92
\end{aligned}
$$

yards. This is the amount of yards player one perceives he will gain. The actual yards gained will be dependent on the actual strategy of player two, this is covered in the post game analysis section. Suppose player one calls a long pass, a safety slips and the offense gains 17 yards. The situation now updates to a new play:

$$\mathbf{S_2} = \{-7, 4 : 45, 45, 1, 10, 12.32, 12.32\}.$$

Player one now takes a more risk averse attitude towards the situation because he has a few more downs to get ten yards and he has crossed mid-field. In the huddle, player one calls his play based on the situation only without perceived information as to the defense of player two. With only the knowledge of all the plays available to

player two,

$$\hat{\gamma}_{\mathbf{S_2}}^{(0)} = \{.0601, .4155, .3129, .2125\} \mid$$

$$\hat{\beta}_{\mathbf{S_2}}^{(0)} = \{4 - 4Overload, 5 - 4Blitz, 4 - 4Zone, 4 - 3Man\}, \quad (23)$$

and the defense calls

$$\hat{\delta}_{\mathbf{S_2}}^{(0)} = \{.0910, .4459, .2414, .2217\}. \quad (24)$$

There is a nice distribution across the offensive plays, running up the middle almost half of the time. As the quarterback approaches the line, he observes that the defense is definitely not in the 4-4 zone,

$$\hat{\beta}_{\mathbf{S_2}}^{(1)} = [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = QBObservation\}]$$

$$= [4 - 4Overload, 5 - 4Blitz, 4 - 3Man].$$

Thus,

$$\hat{\gamma}_{\mathbf{S_2}}^{(1)} = [\gamma | \hat{\beta}_{\mathbf{S_2}}^{(1)}]$$

$$= \{0, 0, .8066, .1934\},$$

running a short pass the majority of the time. This makes sense because the 4-4 zone defends best against the short pass and player one is perceiving this defense to be unavailable to the defense. The perceived strategy of the defense is

$$\hat{\delta}_{\mathbf{S_2}}^{(1)} = \{.8710, 0, .1290, \}.$$

Suppose the coach notices from the press-box that the defense is guarding heavily against the run,

$$
\begin{aligned}
\hat{\beta}_{\mathbf{S_2}}^{(2)} &= [\beta | \boldsymbol{\zeta}^{(2)} = \{\zeta_1 = QBObservation \cap \zeta_2 = CoachObservation\}] \\
&= [4 - 4Overload, 5 - 4Blitz].
\end{aligned}
$$

Now the perceived optimal strategy is

$$
\hat{\gamma}_{\mathbf{S_2}}^{(2)} = \{0, 0, 0, 1\}.
$$

It makes sense that the offense would choose a deep pass in the situation where the defense is guarding heavily against the run. The number of yards the offense expects to gain is

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_2}}^{(2)} &= \hat{\gamma}_{\mathbf{S_2}}^{(2)} \, \hat{\mathbf{R}}_{\mathbf{S_2}}^{(2)} \, \hat{\delta}_{\mathbf{S_2}}^{(2)'} \\
&= [0, 0, 0, 1]
\begin{bmatrix}
-2.7 & 3.8 \\
3.4 & 2 \\
4 & 3.9 \\
6.1 & 7
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix} \\
&= 6.1
\end{aligned}
$$

Suppose the offense ran the long pass, a lineman missed a block, and the quarterback was sacked for a loss of 6 yards. Now,

$$
\mathbf{S_3} = \{-7, 4 : 15, -49, 2, 16, -3.64, -3.64\}.
$$

The initial strategy of player one is

$$
\hat{\gamma}_{\mathbf{S_3}}^{(0)} = \{0, .2854, .3007, .4139\},
$$

and the initial strategy of player two is

$$\hat{\delta}_{\mathbf{S_3}}^{(0)} = \{.3577, 0, .3192, .3231\}.$$

Player one will choose to throw a long or short pass more often in this situation, but still will run up the middle roughly 30% of the time because it is only second down. The defense will not guard against the short run at all because they believe player one to be risk prone, which means player one will not run up the middle. After the quarterback observes the defense to be heavily guarding against the pass,

$$\hat{\beta}_{\mathbf{S_3}}^{(1)} = [4 - 4Zone, 4 - 3Man],$$

the offense calls the play from the updated distribution

$$\hat{\gamma}_{\mathbf{S_3}}^{(1)} = \{0, .8403, 0, .1597\},$$

while the defense is perceived to call

$$\hat{\delta}_{\mathbf{S_3}}^{(1)} = \{.7460, .2537\}.$$

The offense runs up the middle the majority of the time because it observed the defense to be guarding against the pass more heavily. By running this strategy, the offense expects to gain

$$\pi_{\mathbf{S_3}} = 5.0693$$

yards. The offense runs up the middle and gains 11 yards. Thus,

$$\mathbf{S_4} = \{-7, 3 : 57, 40, 3, 5, 29.12, 29.12\},$$

the offense is using a strategy close to the expected case. This results in the initial strategy coming out of the huddle,

$$\hat{\gamma}_{\mathbf{S_4}}^{(0)} = \{.0567, .3975, .3069, .2390\}.$$

The offense either runs or throws a short pass with the highest probability in this situation and can expect to gain

$$\hat{\pi}_{\mathbf{S_4}} = 3.9546$$

yards. The coach observes that the defense is not in 5-4,

$$\begin{aligned}
\hat{\beta}_{\mathbf{S_4}}^{(1)} &= [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = CoachObservation\}] \\
&= [4 - 4Overload, 4 - 4Zone, 4 - 3Man].
\end{aligned}$$

The distribution updates,

$$\hat{\gamma}_{\mathbf{S_4}}^{(1)} = \{0, .5415, .2555, .2030\},$$

as well as the perception of the strategy of player two,

$$\hat{\delta}_{\mathbf{S_4}}^{(1)} = \{.6760, .1357, .1883\}.$$

Because the defense is not using the formation that best guards against the run up the middle, the offense chooses this play with greater probability. The quarterback then observes that the defense is not in any type of zone, thus

$$\begin{aligned}
\hat{\beta}_{\mathbf{S_4}}^{(2)} &= [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = CoachObservation \cap \zeta_2 = QBObservation\}] \\
&= [4 - 4Overload, 4 - 3Man].
\end{aligned}$$

The perceived optimal decision then becomes

$$\hat{\gamma}_{\mathbf{S_4}}^{(2)} = \{0, 0, .7701, .2299\},$$

with the perceived strategy of player two being

$$\hat{\delta}_{\mathbf{S_4}}^{(2)} = \{.8507, .1493\}.$$

The offense calls an audible according the preceding distribution. While running a short pass, the offense gains just 4 yards, not enough for a first down. So,

$$\mathbf{S_5} = \{-7, 3 : 39, 36, 4, 1, 29.12, 29.12\},$$

and the risk strategy remains the same. The initial perceived optimal strategy in this situation is identical to $\hat{\gamma}_{\mathbf{S_4}}^{(0)}$. Upon arrival to the line of scrimmage, the quarterback sees the defense is not in the 4-4 overload, thus

$$
\begin{aligned}
\hat{\beta}_{\mathbf{S_5}}^{(1)} &= [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = QBObservation\}] \\
&= [5 - 4Blitz, 4 - 4Zone, 4 - 3Man],
\end{aligned}
$$

and

$$\hat{\gamma}_{\mathbf{S_5}}^{(1)} = \{.7970, 0, .0101, .1929\}.$$

Since the defense is not perceived to be overloading in the 4-4,

$$\hat{\delta}_{\mathbf{S_5}}^{(1)} = \{.7286, .0563, .2151\},$$

the sweep is called with high probability while still leaving a chance of calling a long pass to keep the defense on their toes. The coach on the sidelines further notices that

the defense is not in a 4-4 of any type so

$$\hat{\beta}_{\mathbf{S_5}}^{(2)} = [\beta | \boldsymbol{\zeta}^{(1)} = \{\zeta_1 = QBObservation \cap \zeta_2 = CoachObservation\}]$$
$$= [5 - 4Blitz, 4 - 3Man],$$

and

$$\hat{\gamma}_{\mathbf{S_5}}^{(2)} = \{0, 0, .7779, .2221\}.$$

The offense abandons the sweep for the short pass as this will yield more yards against the two perceived available defenses. The coach in the pressbox radios down to the head coach exclaiming the defense to be in a 5-4 blitz formation and heavily guarding the run. The perception of the actions available to player two updates

$$\hat{\beta}_{\mathbf{S_5}}^{(3)} = [5 - 4Blitz].$$

The quarterback quickly calls an audible to account for this perception, his optimal strategy being

$$\hat{\gamma}_{\mathbf{S_5}}^{(3)} = \{0, 0, 0, 1\}.$$

The offense calls the long pass, however it is batted down and the defense takes over on downs. This section demonstrated the use of the methodology as applied to a football game. Next, the methods and techniques are presented that allow further optimization of offensive strategies for future situations and corrective action on the strategies used during the game. With this new information, a different outcome may be achieved in future situations.

4.4.4 Post-Game Analysis. During the game, it is best to use design of experiments to explore the interactions between ρ_1 and ρ_2, as this is the most efficient manner to quickly study behavior and outcomes. This is true whenever there are constraints dealing with time or money. In a post game analysis setting with virtually unlimited time, a proper approach would involve exploring each possible situation that
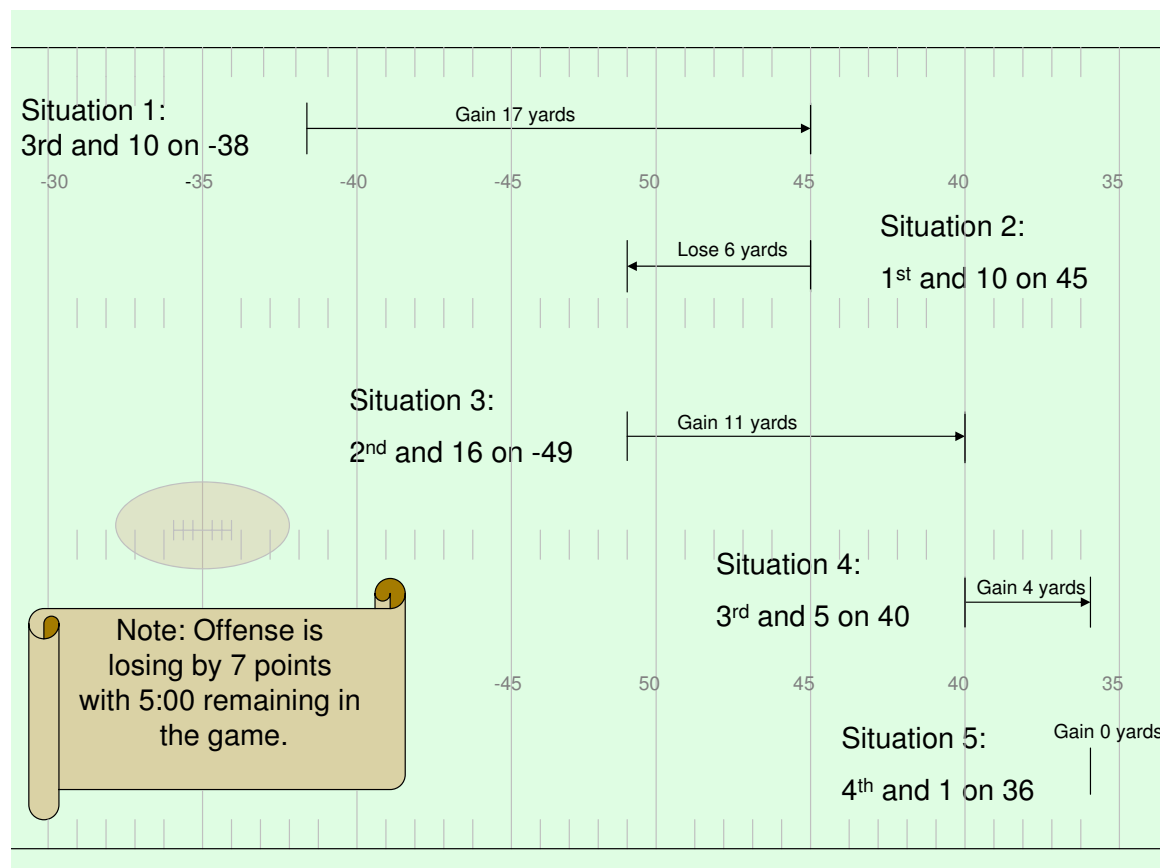
Figure 6:   Football Game Flow Chart

may arise. Following the game, the film could be reviewed and a thorough analysis conducted on each combination of offensive and defensive strategies and utilities. This could easily be done using the techniques presented in the game against nature case. However, this section will demonstrate the use of DOE in the post-game setting. This will lead to learning among the team and possibly the determination of a better strategy to play in future game situations.

Each of the five situations will be examined to determine a better strategy for similar future situations and the quality of observations made by the offense. The interaction plots from the possible risk behaviors of player one and player two can be examined to determine the best strategy for player one. The risk behavior of player two is a noise factor, it cannot be controlled. Therefore, player one may choose to

select the strategy that is most robust to any changes in the risk behavior of player two.

4.4.4.1 Situation 1.   Tending towards a more risk prone risk attitude is the optimal choice when information about the action set available to player two is unavailable. From Figure 7, it appears that the offense was utilizing the best possible strategy in the first situation during the initial game setup. That is, the risk prone approach provides the maximum number of yards gained regardless of the risk strategy of player two. When player two plays the risk averse strategy, player one gains the maximum number of yards by approaching the situation with a risk prone attitude. However, when player two approaches the situation with a risk averse attitude and player one takes a risk averse attitude, player one gains the minimum number of yards. Using the updated information from the quarterback observation, the offense used the poorest possible risk approach with the knowledge it possessed in Equation 20, as shown in Figure 8. The risk prone approach is strictly dominated by both the expected case and the risk averse strategy. If the offense used the expected case, they could have expected to gain almost twice the number of yards as with the risk prone approach.

$$\hat{\gamma}_0^{(1)} = \{0, .9238, 0, .0762\}$$

is the best approach for the offense. The offense is advised to alter its risk strategy in future situations resembling situation 1 where the offense perceives the defense to be in 4-4 zone or a 4-3 man. By playing the risk averse or expected case, the offense guarantees a robust risk approach to the situation.

4.4.4.2 Situation 2.   The initial optimal risk strategy is again that of a risk prone nature as shown in Figure 7. The optimal risk strategy does not change from situation to situation when all the actions are available to player two, however it will differ depending on what action set player one perceives player two is choosing from. From Figure 9, player one appears to have chosen a less than optimal
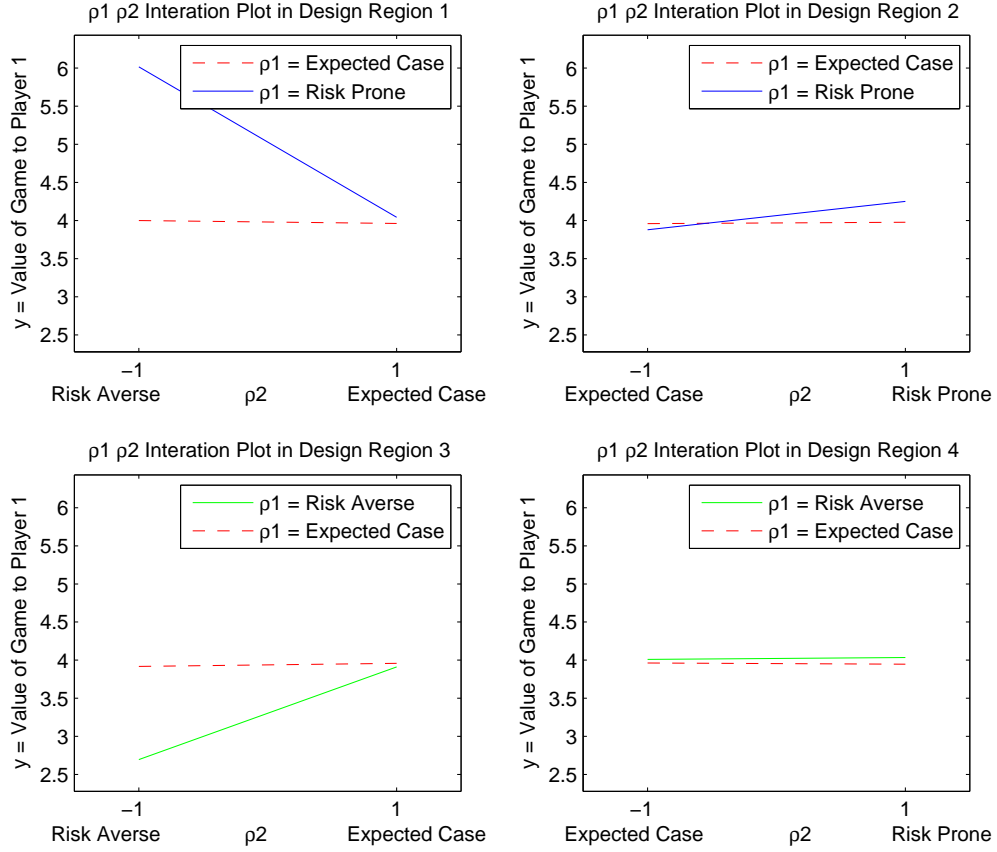
86

Figure 7:   Initial ρ Interaction Plot

risk approach after the update by choosing the more risk averse behavior. When the defense approaches the situation with a risk averse attitude, and the offense also approaches the situation with risk averse attitude, the expected yards gained is only around 4. The offense could have gained the maximum number of yards through the use of a risk prone attitude, however this also introduced the possibility of gaining less yards than the expected case if player two assumed player one to be risk prone. In this situation, since the offense preferred to entertain a risk averse attitude due to its preferences, of equal or better quality risk behavior would have been the expected case. It is robust in that regardless of the risk attitude of player two, player one can expect to gain the same amount of yards.
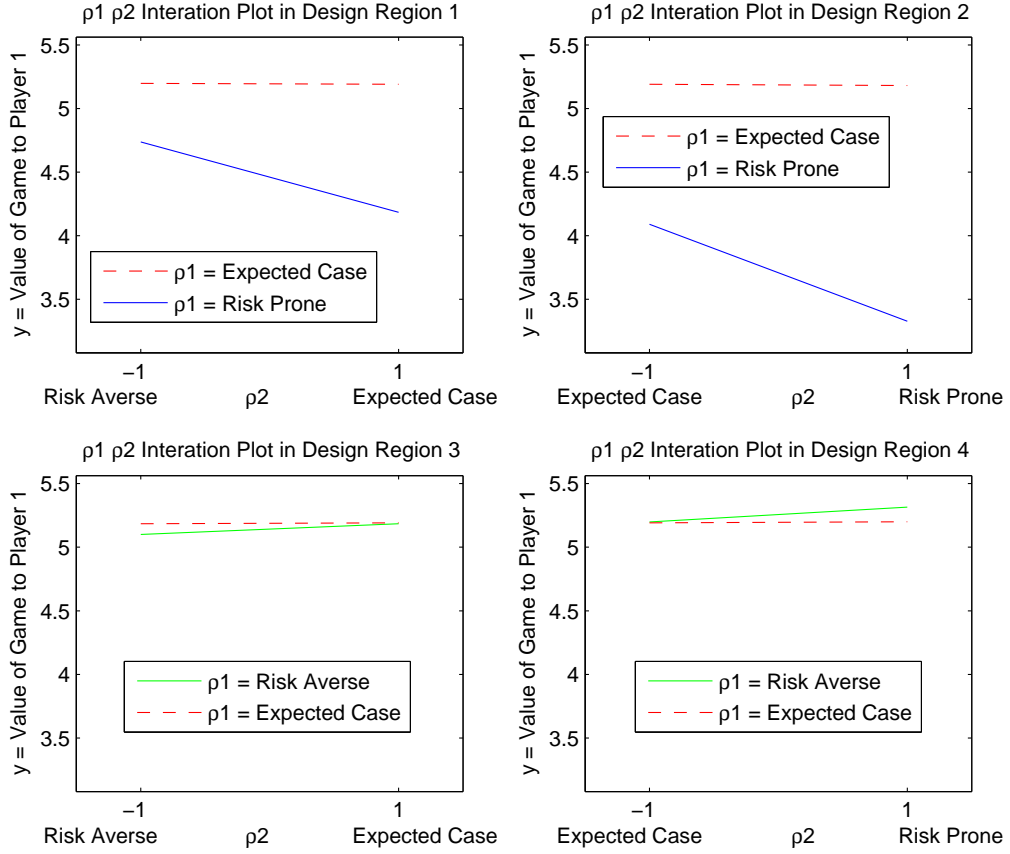
87

Figure 8:   Updated ρ Interaction Plot at S_1

4.4.4.3 Situation 3.   In this situation, the quarterback perceived the defense to again be in a 4-3 man or a 4-4 zone as in situation 1, thus the optimal risk strategy is still available in Figure 8. The risk strategy of the offense in this situation was not as risk prone as in the first situation, leading to a greater probability of calling the run up the middle. The defensive pressure on the pass caused the offense to take a different approach because they were not in an extreme risk behavior situation as in situation 1. The offense was further from the extreme risk prone approach and could expect to gain near the same amount of yards as the expected case during this situation.
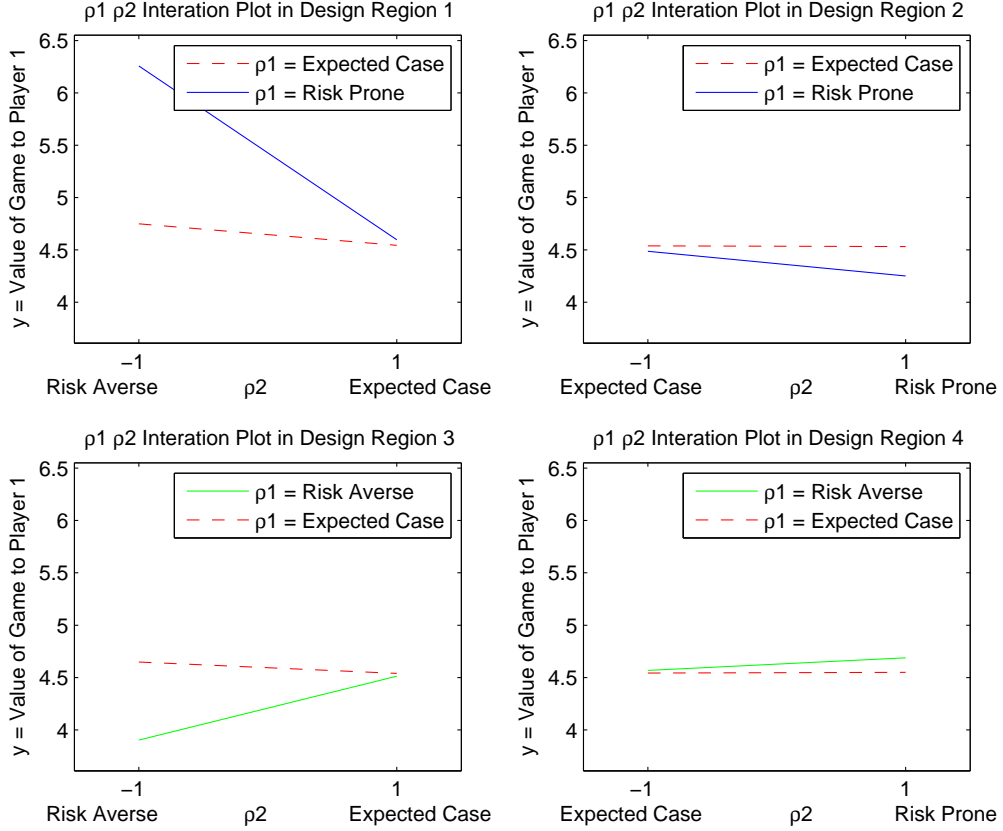
Figure 9:   Updated ρ Interaction Plot at S_2

4.4.4.4 Situation 4.   During situation 4, the first update yields a similar optimal risk strategy approach as that in the initial game setup in Figure 7. After the second update, the interaction plots for the risk strategy are almost identical to that during the second situation at the first update in Figure 9. The offense should have used a more risk prone attitude, possibly gaining more yards than actually achieved during situation 4.

4.4.4.5 Situation 5.   Initially, the best risk approach to take is the same as the above cases, the risk prone strategy. Upon QB observation that the defense is not in a 4-4 overload, the risk prone approach loses value. In Figure 10, when the defense takes the risk prone approach and the offense takes the risk prone
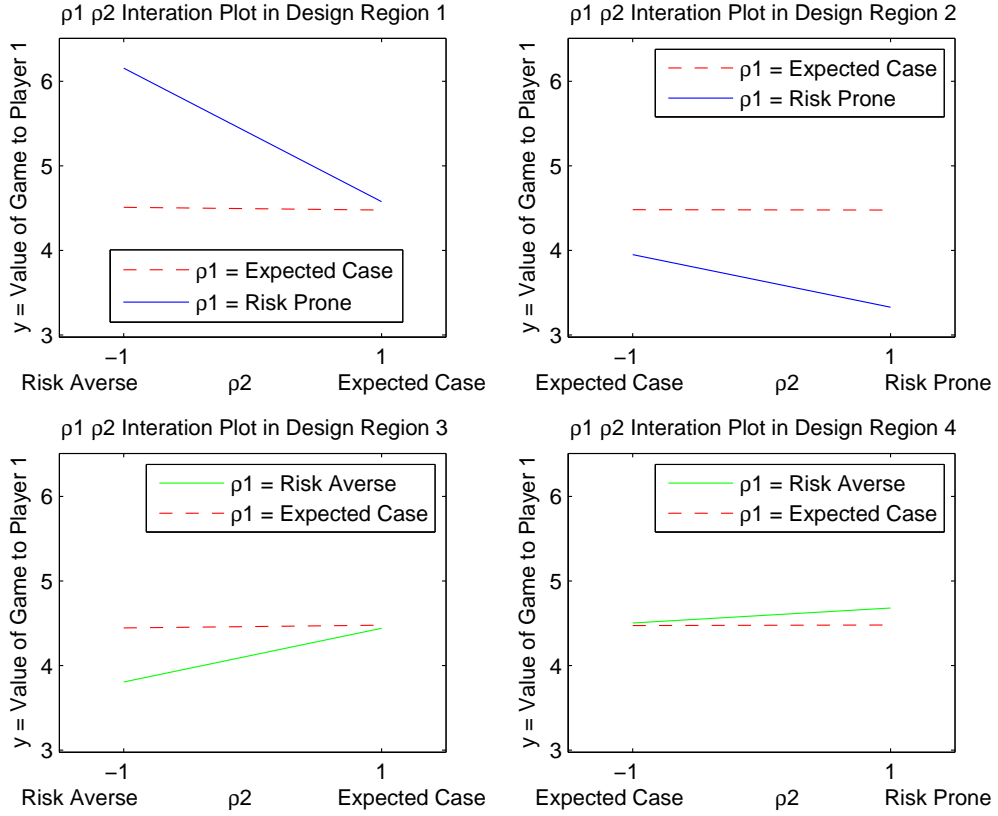
Figure 10:   First Update ρ Interaction Plot at S_5

approach, there is a significant loss. Compared with the robustness of the expected case, the risk prone strategy may not be the best approach. If the defense plays risk averse, and the offense plays risk prone, the offense can expect to gain more yards than in any other situation. The second update results in similar strategy implications as that shown in Figure 11. The risk averse strategy produces similar results, when the defense takes a risk prone approach to the scenario and the offense is risk averse, the offense can expect to gain more yards. When the defense plays risk averse and the offense is risk averse, the offense can expect to gain less yards.

4.4.5 Studying Game Film.   The coaches may be interested in the quality of their observations during games throughout the season. A way to measure how well they are reading the defense is by obtaining the true optimal decisions from past
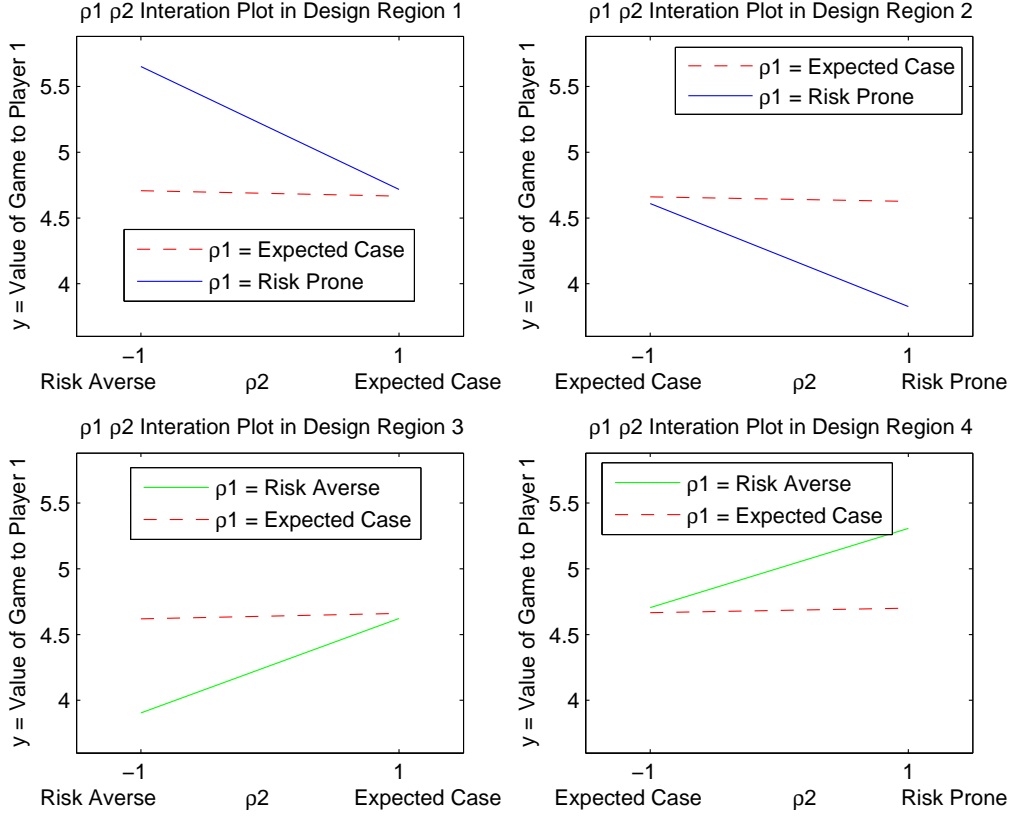
Figure 11: Second Update ρ Interaction Plot at S_5

game tapes and comparing this with the perceived optimal decisions called during the game.

Suppose the true defense in situation 1 was a 4-4 zone. Using Equation 7 on page 26 and the perceived optimal strategy after all the updates, $\hat{\gamma}_{\mathbf{S_1}}^{(1)}$, the value of

the perceived optimal strategy is

$$\hat{\pi}_{\mathbf{S_1}}^{(1)} = \hat{\gamma}_{\mathbf{S_1}}^{(1)} \mathbf{R}_{\mathbf{S_1}}^{(1)} \delta_{\mathbf{S_1}}^{(1)'}$$

$$= [0, 0, .1933, .8067] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1]$$

$$= 5.18,$$

while the true value of the game given the offense knew the defensive formation is

$$\pi_{\mathbf{S_1}}^{(1)} = \gamma_{\mathbf{S_1}}^{(1)} \mathbf{R}_{\mathbf{S_1}}^{(1)} \delta_{\mathbf{S_1}}^{(1)'}$$

$$= [0, 0, 0, 1] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1]$$

$$= 6.3.$$

The true optimal value is calculated using the strategy that the offense would have used had they known the defense was in a 4-4 zone. The difference between these two values is

$$\vec{\pi}_{\mathbf{S_1}}^{(1)} = \pi_{\mathbf{S_1}}^{(1)} - \hat{\pi}_{\mathbf{S_1}}^{(1)}$$

$$= 5.18 - 6.3$$

$$= -1.12$$

yards. This is the lost opportunity by the offense for not having perfect information. That is, the offense lost 1.12 yards because they could only perceive the defensive formation in part. To determine the value of the quarterback observation in this scenario, the value of the game at $s = 0$ must be subtracted from the value at $s = 1$.

The value at $s = 0$, or the original value of the game taking into consideration the risk strategy of the players $\rho_1 = \rho_2 = -.7$, is

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_1}}^{(0)} &= \hat{\gamma}_{\mathbf{S_1}}^{(0)} \, \mathbf{R}_{\mathbf{S_1}}^{(0)} \, \delta_{\mathbf{S_1}}^{(0)'} \\
&= [0, 0, .1538, .8462] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} \quad [1] \\
&= 5.41.
\end{aligned}
$$

Thus the value added by the quarterback is

$$
\begin{aligned}
\dot{\pi}_{\mathbf{S_1}}^{(1)} &= \hat{\pi}_{\mathbf{S_1}}^{(1)} - \hat{\pi}_{\mathbf{S_1}}^{(0)} \\
&= 5.18 - 5.41 \\
&= -.23.
\end{aligned}
$$

The quarterback observation in this situation, even though correct, actually cost the offense .23 yards of expected gain. Normally, a good observation will add value to the game, this is a rare exception.

During situation 2, suppose the defense was truly in a 5-4 Blitz formation. In this situation, the offense chose to run the deep pass with probability 1 because

$$
\hat{\beta}_{\mathbf{S_2}}^{(2)} = [4 - 4Overload, 5 - 4Blitz].
$$

In this situation, even though the offense did not have full information about the defensive formation, his strategy was such that

$$
\pi_{\mathbf{S_2}}^{(2)} = \hat{\pi}_{\mathbf{S_2}}^{(2)} = 7.
$$

The true value of the game given the 5-4 Blitz is equal to the perceived value of the game given the 5-4 Blitz as calculated in Equation 25. Gaining perfect information in this situation would not be of value to the offense. The original strategy by the offense, $\hat{\gamma}^{(0)}$, results in

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_2}}^{(0)} &= \hat{\gamma}_{\mathbf{S_2}}^{(0)} \, \mathbf{R}_{\mathbf{S_2}}^{(0)} \, \delta_{\mathbf{S_2}}^{(0)'} \\
&= [.0601, .4155, .3120, .2125] \begin{bmatrix} 3.8 \\ 3 \\ 3.9 \\ 7 \end{bmatrix} [1] \\
&= 3.76.
\end{aligned}
$$

The first observation by the quarterback results in a value of

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_2}}^{(1)} &= \hat{\gamma}_{\mathbf{S_2}}^{(1)} \, \mathbf{R}_{\mathbf{S_2}}^{(1)} \, \delta_{\mathbf{S_2}}^{(1)} \\
&= [0, 0, .8066, .1934] \begin{bmatrix} 3.8 \\ 3 \\ 3.9 \\ 7 \end{bmatrix} [1] \\
&= 4.50.
\end{aligned}
$$

Thus the value added by the quarterback is

$$
\begin{aligned}
\dot{\pi}_{\mathbf{S_2}}^{(1)} &= \hat{\pi}_{\mathbf{S_2}}^{(1)} - \hat{\pi}_{\mathbf{S_2}}^{(0)} \\
&= 4.5 - 3.76 \\
&= .74.
\end{aligned}
$$

The quarterback observation in this situation adds about .75 yards of expected gain by the offense, this is good. The value of the coaches observation is found using the

value of the game at s=2

$$\hat{\pi}_{\mathbf{S_2}}^{(2)} = \hat{\gamma}_{\mathbf{S_2}}^{(2)} \mathbf{R}_{\mathbf{S_2}}^{(2)} \delta_{\mathbf{S_2}}^{(2)'}$$

$$= [0, 0, 0, 1] \begin{bmatrix} 3.8 \\ 3 \\ 3.9 \\ 7 \end{bmatrix} [1] \tag{25}$$

$$= 7.$$

The added value of the coaches observation is

$$\dot{\pi}_{\mathbf{S_2}}^{(2)} = \hat{\pi}_{\mathbf{S_2}}^{(2)} - \hat{\pi}_{\mathbf{S_2}}^{(1)}$$

$$= 7 - 4.5$$

$$= 2.5$$

yards.

Situation 3 finds the defense truly in the 4-3 Man formation. Recall the offense perceived the actions available to the defense as $\hat{\beta}_{\mathbf{S_3}}^{(1)} = [4 - 4Zone, 4 - 3Man]$. The value of the perceived optimal strategy is thus

$$\hat{\pi}_{\mathbf{S_3}}^{(1)} = \hat{\gamma}_{\mathbf{S_3}}^{(1)} \mathbf{R}_{\mathbf{S_3}}^{(1)} \delta_{\mathbf{S_3}}^{(1)'}$$

$$= [0, .8403, 0, .1597] \begin{bmatrix} 6.6 \\ 5.9 \\ 7.3 \\ -3.4 \end{bmatrix} [1]$$

$$= 4.41,$$

while the true value of the game is

$$
\begin{aligned}
\pi_{\mathbf{S_3}}^{(1)} &= \gamma_{\mathbf{S_3}}^{(1)} \, \mathbf{R}_{\mathbf{S_3}}^{(1)} \, \delta_{\mathbf{S_3}}^{(1)'} \\
&= [0, 0, 1, 0] \begin{bmatrix} 6.6 \\ 5.9 \\ 7.3 \\ -3.4 \end{bmatrix} [1] \\
&= 7.3.
\end{aligned}
$$

The difference

$$
\begin{aligned}
\vec{\pi}_{\mathbf{S_3}}^{(1)} &= \pi_{\mathbf{S_3}}^{(1)} - \hat{\pi}_{\mathbf{S_3}}^{(1)} \\
&= 7.3 - 4.41 \\
&= 2.89
\end{aligned}
$$

yards represents the number of yards the offense could expect to gain had they known the defense was in a 4-3 Man. The value added by the quarterback observation is calculated as in the previous 2 situations.

$$
\begin{aligned}
\dot{\pi}_{\mathbf{S_3}}^{(1)} &= \hat{\pi}_{\mathbf{S_3}}^{(1)} - \hat{\pi}_{\mathbf{S_3}}^{(0)} \\
&= 4.41 - 2.47 \\
&= 1.94,
\end{aligned}
$$

is the number of yards gained due to the quarterback observation.

During situation 4, the offensive strategy failed to include the true action of the defense as a possible action, the defense was actually in a 4-4 zone. The offense chose to throw the short pass in the situation and only gained a few yards. The perceived optimal decision was based on

$$
\hat{\beta}_{\mathbf{S_4}}^{(2)} = [4 - 4Overload, 4 - 3Man].
$$

96

This resulted in a low value of the perceived optimal decision,

$$\hat{\pi}_{\mathbf{S_4}}^{(2)} = \hat{\gamma}_{\mathbf{S_4}}^{(2)} * \mathbf{R}_{\mathbf{S_4}}^{(2)} \delta_{\mathbf{S_4}}^{(2)'}$$

$$= [0, 0, .7701, .2299] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1]$$

$$= 1.83,$$

while the true value of the game is

$$\pi_{\mathbf{S_4}}^{(2)} = \gamma_{\mathbf{S_4}}^{(2)} \mathbf{R}_{\mathbf{S_4}}^{(2)} \delta_{\mathbf{S_4}}^{(2)'}$$

$$= [0, 0, 0, 1] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1]$$

$$= 6.3.$$

The difference

$$\vec{\pi}_{\mathbf{S_4}}^{(2)} = \pi_{\mathbf{S_4}}^{(2)} - \hat{\pi}_{\mathbf{S_4}}^{(2)}$$

$$= 6.3 - 1.83$$

$$= 4.47$$

yards is significant and probably would have gained the offense a first down had they properly perceived the situation even to some degree. The yards gained by the original

coach observation was

$$\dot{\pi}_{\mathbf{S_4}}^{(1)} = \hat{\pi}_{\mathbf{S_4}}^{(1)} - \hat{\pi}_{\mathbf{S_4}}^{(0)}$$
$$= 4.17 - 3.91$$
$$= .26,$$

while the added yards of the QB observation was

$$\dot{\pi}_{\mathbf{S_4}}^{(2)} = \hat{\pi}_{\mathbf{S_4}}^{(2)} - \hat{\pi}_{\mathbf{S_4}}^{(1)}$$
$$= 1.83 - 4.17$$
$$= -2.34.$$

Individually studying the value added by the observations leads us to conclude the QB observation was in error and cost the offense about 2.34 yards.

During the final situation, the offense perceived the defense to be in 5-4 Blitz formation and decided to call a deep pass. The ball was knocked down and the defense took over on downs. Recall that

$$\hat{\beta}_{\mathbf{S_5}}^{(2)} = [5 - 4Blitz, 4 - 3Man],$$

was correct up to the second update. The true defensive formation was a 4-3 Man, the observation from the press box was in error,

$$\vec{\pi}_{\mathbf{S_5}}^{(3)} = (-3.4) - 7.3$$
$$= -10.7.$$

The strategy used by the offense resulted in a loss of opportunity of 10.7 yards. With true information, the offense would have surely gained a first down and possibly won the game.

The performance of the offense can be graphed over time to determine the quality of the reads by the quarterback, sideline coach, and pressbox coach collectively and individually. Using the value of perfect information $\vec{\pi}$ at each situation gives a feel for the performance. The difference between the true and perceived optimal decisions can also be thought of as the regret or error that the offense shouldered because of their perception. If the perception by the offense is the defensive formation that the defense actually runs, perfect information has no value. When the offensive perception is in error or incomplete, the offensive regret increases, or the value of obtaining perfect information increases. A high value of perfect information indicates a poor perception by the offense. Looking at Figure 12, the error of the offensive strategy increased over
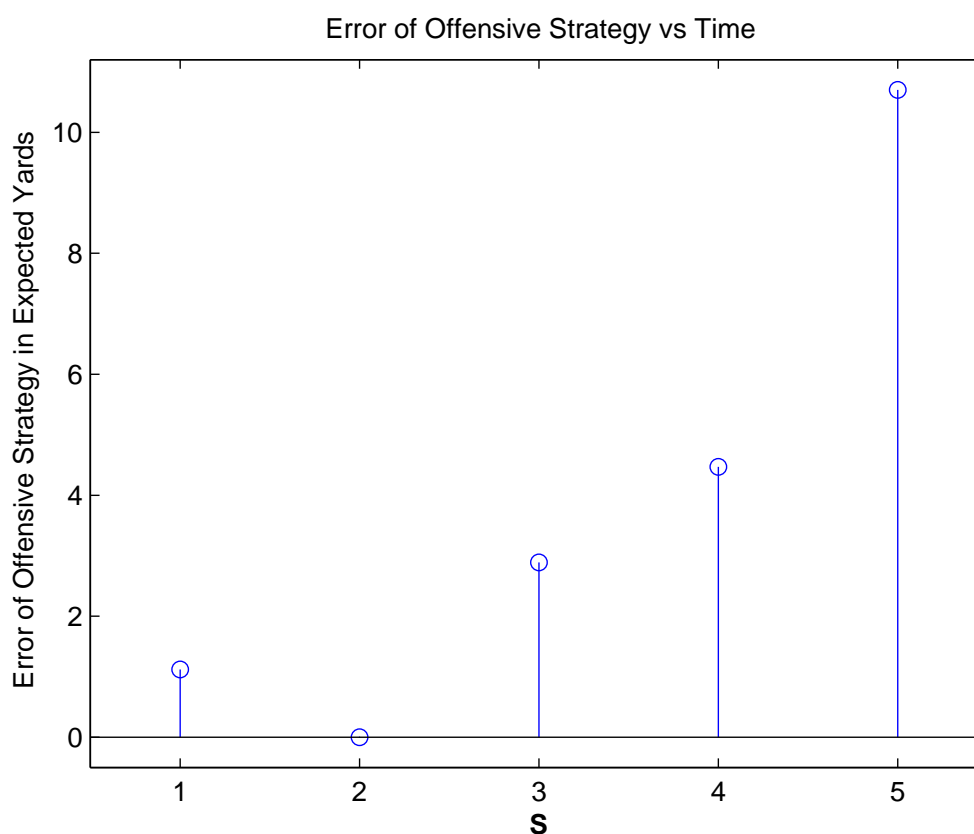

Figure 12:   Value of Perfect Information

time, indicating that the offensive perception of the defensive formation degraded as the game progressed. This is computed by using the true optimal strategy given the

truth in each situation. This is subtracted from the value of the perceived optimal strategy given the truth. For example, during situation 1, the true defensive formation was the 4-4 zone. The true optimal offensive strategy for the 4-4 zone is

$$\gamma_{\mathbf{S_1}} = \{0, 0, 0, 1\}.$$

This results in a true optimal value of the game

$$
\begin{aligned}
\pi_{\mathbf{S_1}} &= \gamma_{\mathbf{S_1}} \, \mathbf{R_{S_1}} \, \delta'_{\mathbf{S_1}} \\
&= [0, 0, 0, 1] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1] \\
&= 6.3.
\end{aligned}
$$

The perceived value of the game is

$$
\begin{aligned}
\hat{\pi}_{\mathbf{S_1}} &= \hat{\gamma}_{\mathbf{S_1}} \, \mathbf{R_{S_1}} \, \delta'_{\mathbf{S_1}} \\
&= [0, 0, .1933, .8067] \begin{bmatrix} 4 \\ 5.1 \\ .5 \\ 6.3 \end{bmatrix} [1] \\
&= 5.18.
\end{aligned}
$$

Subtracting these two values gives the error of the offensive strategy during situation 1, $6.3 - 5.18 = 1.12$. The performance of each individual observer can also be examined over time. The value added in relation to the previous time step or update can be determined. Figure 13, shows the value added to the game by the quarterback over time. The observation of the quarterback hurt the value of the game during situation 4. The value of all the observations is a better measure of how well the
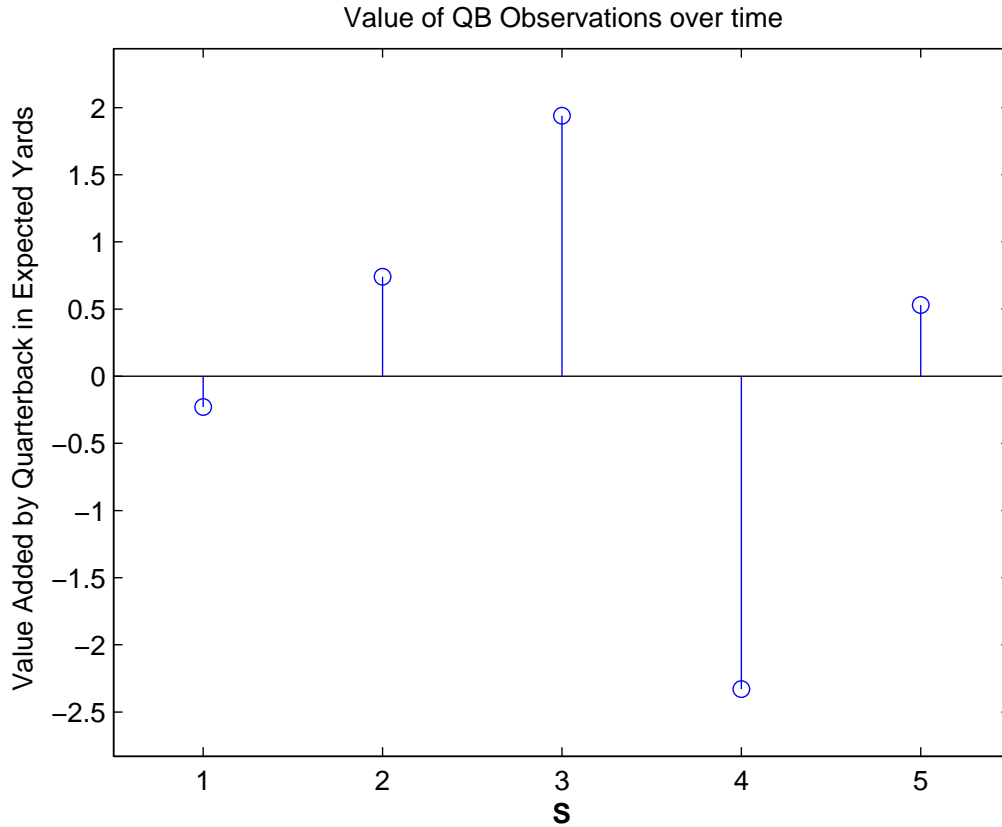
100

Figure 13:   QB Added Value over time

offense is reading the defense, as the observations or sensors are not independent, each observation relies on the previous. Figure 14 shows the value of all the observations over time. During situations 2 and 3, the observations of the offense added a significant number of yards to the expected gain, while during situations 4 and 5, the offensive observations actually impaired their strategy.

These are simple examples, in reality this graph may contain much more insight into the performance of an individual over the course of a season or game. This may indicate things such as fatigue during a game, or learning during the game. Over a season, these graphs could show the maturity gained by a junior quarterback, or lack there of. In this way, the performance of individuals or coaches (sensors) can be analyzed.
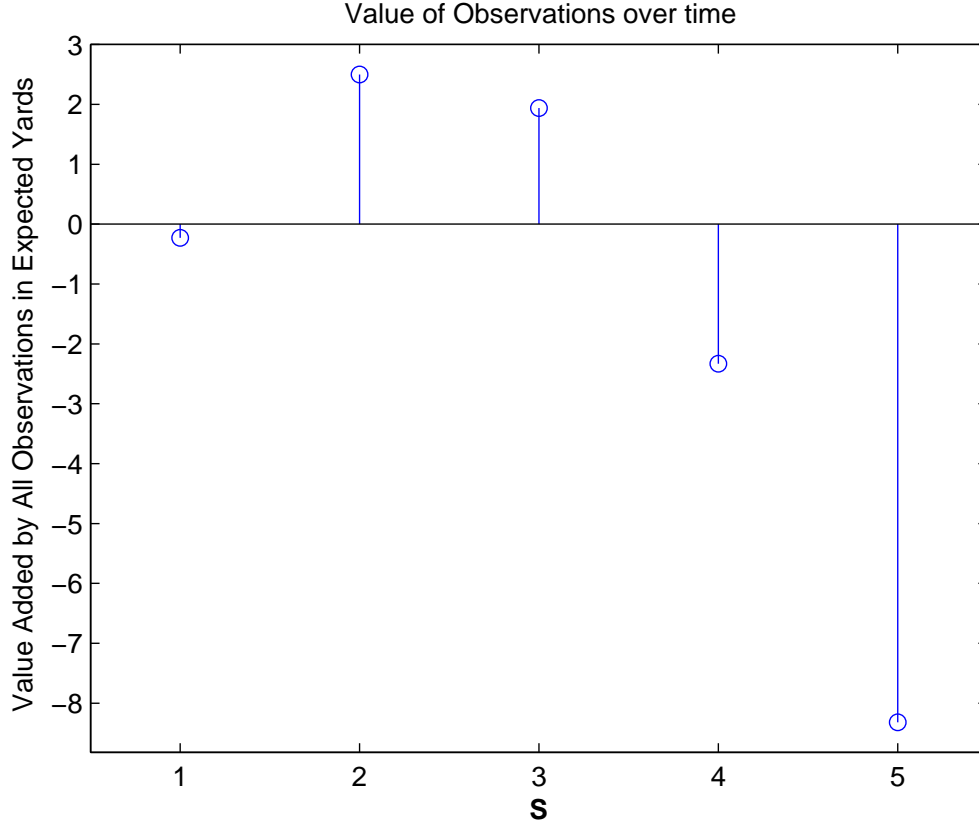
Figure 14:   Observations Added Value over time

4.5   Allocations of Financial Funds

4.5.1   Introduction.   This methodology can be used in various financial situations. A major beneficiary of this research could be the government. Military and government budgeting is certainly dependent on the many given states of the world. As new information is gained over time, resources need to be optimally dispersed to ensure they are utilized properly and the contributions of the citizens are not squandered in needless pursuits. An area of high visibility at the present age is the proper way to allocate our nations resources in defense against potential terrorist attacks. This can be modeled as a two-player game where player one is the Department of Homeland Security (DHS) and player two is the terrorist regime. The action sets of player one and player two are the amount of resources to allocate to certain eco-

nomic areas and possible targets to attack, respectively. As intelligence information is received about the desirable attack locations of the terrorists, our resources can be properly allocated to reduce the damage that will result from the attacks. This section will present a simple example of how this methodology could be used to reduce the amount of damage caused by terrorists.

4.5.2 Resource Allocations of Terrorist Funds. There has recently been much research done in the area of reward matrices for possible terrorist attacks and the amount of resources allocated to the particular target. For this example, the reward matrix is kept simple, used solely for the purposes of demonstrating the methodology. The reward matrix in this example will assume the same scale used in the combat game example, a Likert scale from -5 to 5, with -5 being the worst possible outcome. After speaking with the DHS, the outcomes from the experts are given in Table 28. Keep in mind, the values here are just rated values. In reality, these will be some

Table 28: Normal Form of Terrorist Resource Allocation

	Airline	Subway	Downtown Businesses	Anthrax Mail
Public Transportation	5	4.5	-2	-5
Government Agencies	-5	-4	-1	3
Urban Areas	-3	-2	5	-4

function of lives lost, resources, and economic impact for example.

Suppose that information is unavailable about the intent of the terrorist organization, each of the four targets are possible areas for attack. The strategy of player one is calculated as in the previous examples,

$$\hat{\gamma}_{\rho_1}^{(s)} = \{w_1, w_2, w_3\}$$
$$\hat{\gamma}_0^{(0)} = \{.3977, .5088, .0936\}$$

these being the percentages of resources to allocate to the three areas of protection. The strategy of the terrorists is

$$\hat{\delta}_0^{(0)} = \{.3216, 0, .3158, .3626\}.$$

The consequences of different risk behavior are shown in Figure 15. The risk prone
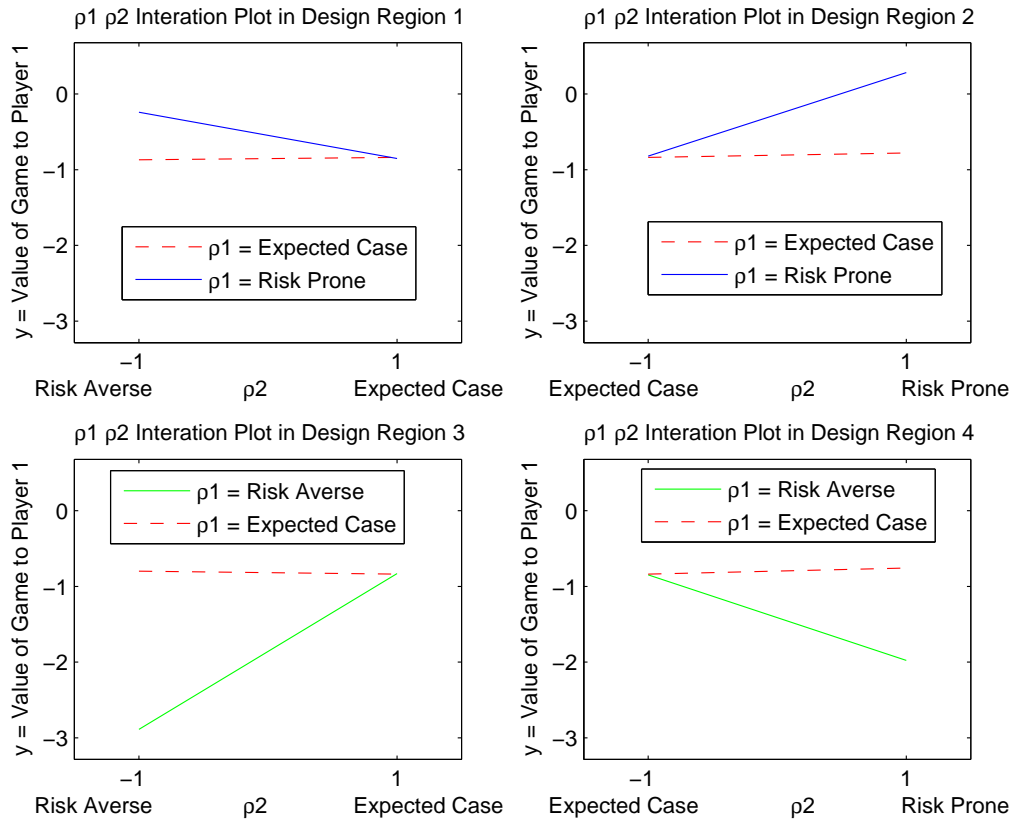

Figure 15:   Initial Effects of Risk Behavior

strategy appears to be the best strategy for the DHS without any information about the actions of the terrorists. Regardless of the risk behavior of the terrorists, the risk prone strategy gives higher payoffs. The best perceived allocation of resources is then

$$\hat{\gamma}_{-1}^{(0)} = \{.1644, .7378, .0979\}.$$

Notice this still results in a loss the majority of the time, but is the best strategy player one can play in this situation.

Next, suppose the DHS received intelligence that the terrorists were abandoning attacks on the public transportation, the airlines and the subway, because the security measures imposed by the U.S. had increased the difficulty to a level far too great for the terrorists to achieve results. Now,

$$
\begin{aligned}
\hat{\beta}^{(1)} &= [\beta | \zeta^{(1)} = \{\zeta_1 = Intelligence\}] \\
&= [DowntownBusinesses, AnthraxMail]
\end{aligned}
$$

The strategy of player one will update due to his perception of the action set of the terrorists, that is

$$
\begin{aligned}
\hat{\gamma}_0^{(1)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{0, .6923, .3077\}.
\end{aligned}
$$

It makes sense that the DHS would remove funding from an area where no threat was present. Again, this example is extreme and is for demonstration purposes only. Furthermore, the perceived probabilities of attacks are

$$
\hat{\delta}_0^{(1)} = \{.5385, .4615\}.
$$

The effects of risk behavior in this situation can be seen in Figure 16. In this case, taking the risk prone or risk averse attitude towards the situation could result in great gains OR great losses. The expected case is the most robust risk strategy to approach the situation with. Regardless of the risk strategy of the terrorists, the DHS is guaranteed a gain of around 1. This allocation of resources will protect best against the terrorists regardless of their approach to attacks.
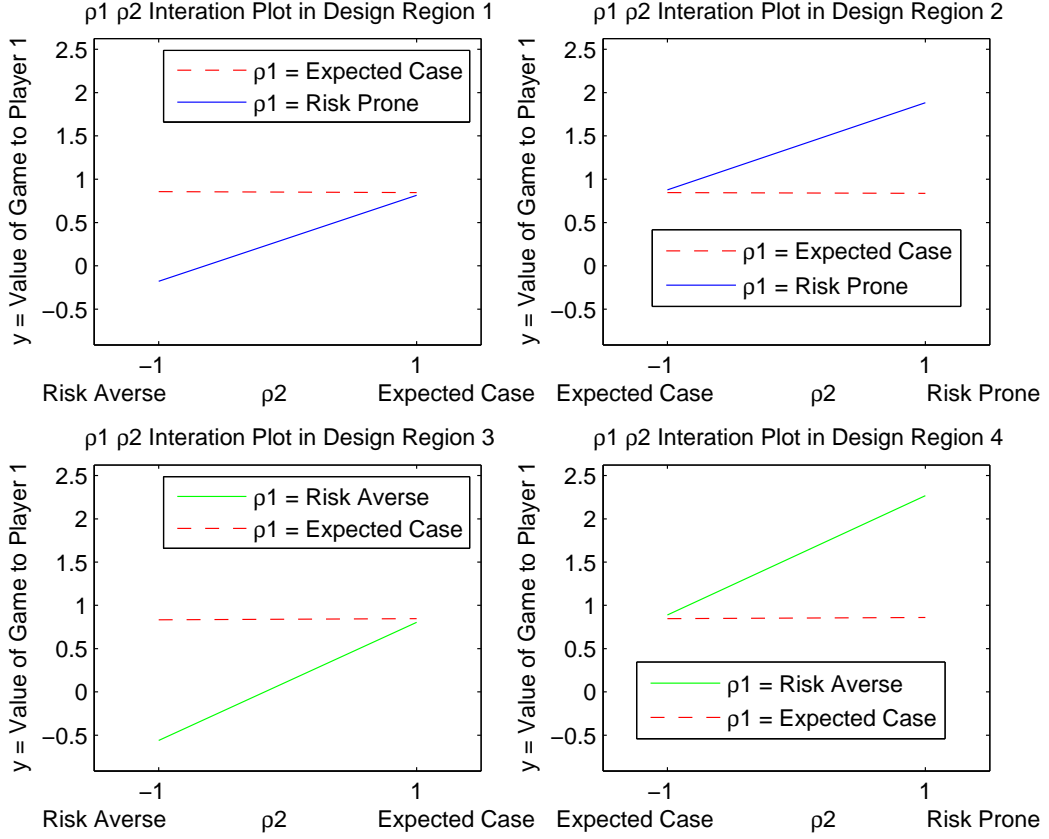
Figure 16:   Updated Effects of Risk Behavior

Suppose finally that the DHS received information from a CIA spy that the terrorists had ceased talk about attacking downtown businesses and increased talks about attacks on public transportation, subways and airports. Thus,

$$
\begin{aligned}
\hat{\beta}^{(2)} &= [\beta | \zeta^{(2)} = \{\zeta_1 = Intelligence \cap \zeta_2 = CIA\ Spy\}] \\
&= \{Airline, Subway, AnthraxMail\},
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{\gamma}_0^{(2)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{.4444, .5556, 0\}.
\end{aligned}
$$

106

The perceived terrorist strategy is

$$\hat{\delta}_0^{(2)} = \{.1644, .7378, .0979\},$$

with a resulting perceived value of

$$
\begin{aligned}
\hat{\pi}_{0(0)}^{(2)} &= \hat{\gamma}_0^{(2)} \, \mathbf{R}^{(2)} \, \delta_{\mathbf{0}}^{(2)'} \\
&= [.4444, .5556, 0]
\begin{bmatrix}
5 & 4.5 & -5 \\
-5 & -4 & 3 \\
-3 & -2 & -4
\end{bmatrix}
\begin{bmatrix}
.1644 \\
.7378 \\
.0979
\end{bmatrix} \\
&= -.5556.
\end{aligned}
$$

The risk behavior interactions can be seen in Figure 17. The DHS does not want to be risk averse in this situation as this will lead to the greatest loss regardless of the actions of the terrorists. The risk prone approach is clearly the best risk strategy,

$$
\begin{aligned}
\hat{\gamma}_{-1}^{(2)} &= [\gamma | \hat{\beta}^{(1)}] \\
&= \{.1823, .8177, 0\}.
\end{aligned}
$$

This results in a value of

$$
\begin{aligned}
\hat{\pi}_{-1(0)}^{(2)} &= \hat{\gamma}_{-1}^{(2)} \, \mathbf{R}^{(2)} \, \delta_{\mathbf{0}}^{(2)'} \\
&= [.1823, .8177, 0]
\begin{bmatrix}
5 & 4.5 & -5 \\
-5 & -4 & 3 \\
-3 & -2 & -4
\end{bmatrix}
\begin{bmatrix}
.1644 \\
.7378 \\
.0979
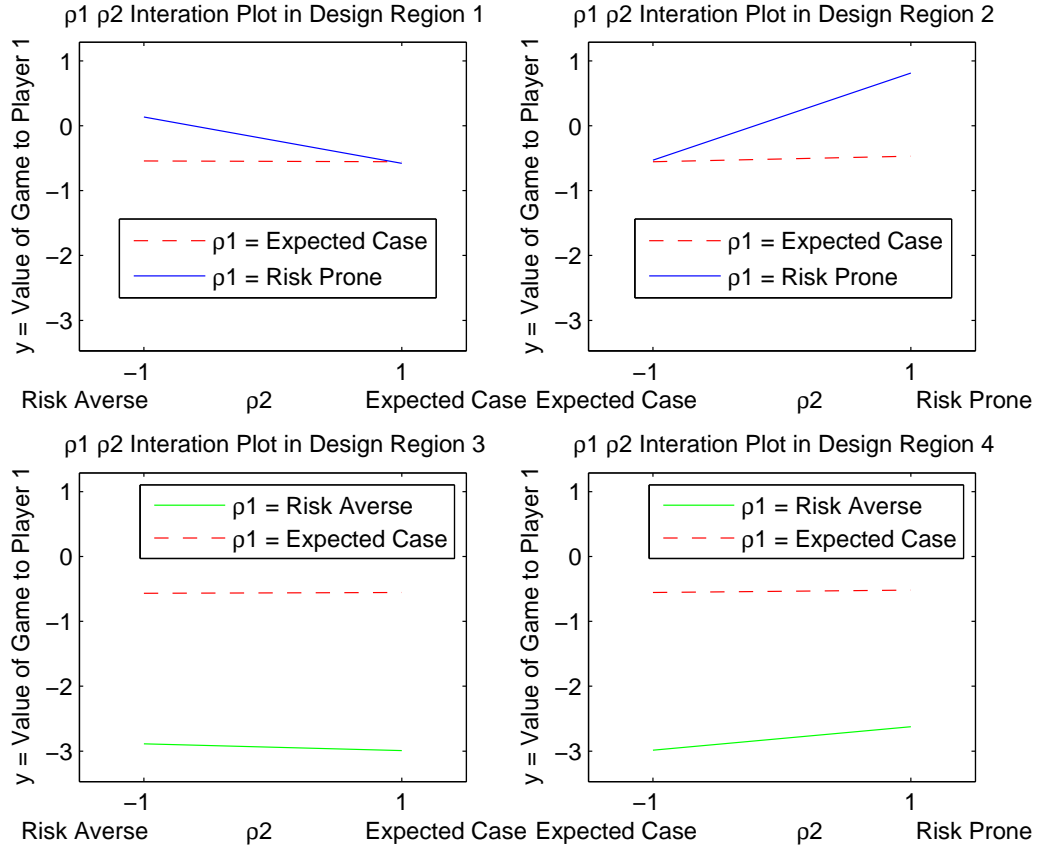\end{bmatrix} \\
&= -.5553.
\end{aligned}
$$

Figure 17: Updated Effects of Risk Behavior

During the holiday season, the terrorists may be more risk prone in their approach, and thus the value of the game could be

$$
\begin{aligned}
\hat{\pi}^{(2)}_{-1(-1)} &= \hat{\gamma}^{(2)}_{-1} \, \mathbf{R^{(2)}} \, \delta^{(2)'}_{-1} \\
&= [.1823, .8177, 0] \begin{bmatrix} 5 & 4.5 & -5 \\ -5 & -4 & 3 \\ -3 & -2 & -4 \end{bmatrix} \begin{bmatrix} .1644 \\ .7378 \\ .0979 \end{bmatrix} \\
&= .8135,
\end{aligned}
$$

if the terrorists took an extreme risk approach. By playing risk prone, we protect ourselves best against a risk prone strategy by the terrorists.

108

In conclusion, the terrorists know that public transportation is very important to us, thus they will actually attack it less. Since we both know the importance of it, it is actually played less because of the dynamics of game theory. This section shows a very simple yet demonstrative use of the methodology on a resource allocation problem. The adversarial nature of the terrorists results in a natural application of game theory.

4.6 Conclusion

This chapter presented the results of using the methodology from chapter 3 to update optimal decisions and measure differences between perceived optimal and optimal decisions. The results are entirely intuitive and show that the methodology could surely be used to accurately represent situations in a simulation model as well as to make actual decisions based on information available. The next chapter presents a conclusion of the work accomplished and the direction for future research.

V. Conclusion and Recommendations

5.1 Introduction

This thesis presents a methodology for updating optimal decisions over time as new information is obtained. The three objectives of the research are:

1. Develop a methodology that automatically updates an optimal decision over time based on the information available to a decision maker at each time step

2. Develop the methodology to capture the effects of incomplete or inaccurate information by measuring the difference between the perceived optimal decision that is based on this inaccurate information and the true optimal decision which is based on perfect information

3. Present a technique to explore the implications of decision maker risk behavior and subsequently suggest better alternatives

The first objective was accomplished using game theory. Methodology was developed that allows an optimal decision to be updated based on the perceived actions available to the other players of the game.

The second objective was completed using the methodology presented in the first objective and further developing the methodology to capture this difference between the perceived optimal decision and the true optimal decision using the value of the game.

The third objective was accomplished through the use of utility theory and response surface methodology. Utility theory is used to transform the reward matrices to produce different strategies for different types of players. A good risk strategy to approach a situation with is then determined subject to the amount of variability in the value of the game willing to be accepted, this variability being completely explained by ρ. This is done through exploring the surface of the response, the value of the game, and is different for each game encountered.

5.2 Model Assumptions

Many of the assumptions considered while using the presented techniques could be eliminated through further research. This initial study provides the framework for updating optimal decisions using game theory in zero-sum, static two-player games with complete information. The main assumptions underlying the game theoretic approach used in this research are

1. Minimax/Maximin Principle - The players of the game are rational decision makers. Player one is trying to maximize his minimum gain while player two is trying to minimize the maximum gain of player one.

2. Zero-sum - The rewards of the outcome sum to zero. The gain of player one is the same as the loss of player two.

3. Sequential and Simultaneous - This theory is sequential in that each player makes decisions based on the perception of the available actions to the other player. However, it is simultaneous in that each player makes a decision without knowing the moves of the other player with certainty.

4. Non-Cooperative - The players of the game are in a conflict with one another and the chance for cooperative bargaining to arise is zero.

5. Static Rewards - The rewards of the players of the game do not change over time.

6. Complete Information - Each player knows the reward matrix with certainty.

5.2.1 Model Strengths. There are numerous strengths in using these techniques. Specifically, it allows a decision maker to update the optimal decision policy based on new information as it arrives. Utility theory allows flexibility in this model by allowing any type of decision maker to be represented. Exploring good risk strategies in approach to each situation further strengthens the quality of the decision. This methodology has many strengths and assumptions which opens the door for future research.

5.2.2 Alternative Application Areas. This study focused on combat and sports games to demonstrate the usefulness of the methodology. Clearly, this research could be applied to a plethora of research areas. Determining the proper allocation of resources as new information becomes available would be useful at the personal and corporate levels. Presently, an entire re-formulation of the problem must occur to update the proper allocations. This research could allow an efficient update based only on new information about nature or the moves of other companies as it is perceived.

Considering the adversarial nature of terrorists, this methodology could be used to determine optimal allocation of resources in defense of our nations assets based on new information as it becomes available. Each time we receive intelligence about the actions of terrorist groups, our optimal allocation of resources will update. The utility of the reward matrix will also change over time for the terrorists. For instance, during the holiday season, the utility of a successful attack on the airports is higher for the terrorist. The usefulness of these methods on this problem was demonstrated in chapter 4, however more research needs to be done to accurately capture the context of the game.

Certainly, this research can be applied in the manufacturing arena. As demand is observed over time, how can supply be optimally updated? Also, the amount of risk a company is willing to take to achieve greater gains becomes of special importance.

Furthermore, this theory can be used in other sports application areas, providing teams with the best strategy to use based on the information they are observing about the behavior of the opponents.

5.3 Further Research

This study generates copious follow on research opportunities. This is the first use of game theory to update optimal decisions, thus the various directions of game theory not touched during this research are ripe for immediate attention. These include multi-player games, dynamic reward matrices, non-zero sum games where

each player has different rewards, cooperative games where players consider alliances, and incomplete information games. These are just a few of the many areas that need expanding after the release of this research.

Much research has been accomplished with regard to the proper specification of the reward matrix. These accurate reward matrices need to be applied to this research.

Chapter four presented an example using a priori distributions to base an optimal decision policy on in the game against nature case. This needs to be formalized as normally the decision maker will have information about the other player of the game. Also, this needs to be expanded to the player one versus player two game to account for prior knowledge about the actions of the other decision maker. This will account for the assumption that player two is always attempting to minimize the maximum loss of player one.

A continuous scale needs to be employed for the automated risk tolerance design matrix, so any values can be plugged in for time, down, distance, etc. This research only considers the discrete case. Additionally, the automation of the risk preference was determined through questioning individual decision makers. The determination of the appropriate function of the input factors and the risk preference was different for each decision maker. There is probably a general model that will accurately describe every type of decision maker, at least approximately for each type of game. This deserves future research, as it would lead to a more efficient approximation of the risk tolerance than the method presented.

```matlab
function [Player1Strategy,Player2Strategy] = OptimalStrategy(R)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%{

Function:

    [Player1Strategy,Player2Strategy] = OptimalStrategy(R)

Author:

    Jeremy D. Jordan, Capt, USAF

Description:

    Calculates the optimal strategies of the players of a two-player
    game.

Inputs:

    R: The original reward matrix the strategies are calculated from

Outputs:

    Player1Strategy: The optimal strategy of player one given the
    reward matrix.

    Player2Strategy: The optimal strategy of player two given the
    reward matrix

%}
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Player Strategies using Game Theory Algorithm %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%% Calculate Player 1 and Player 2 optimal strategies %%%%

%% Size of Reward Matrix
s=size(R);

%% Number of rows in Reward Matrix
```

```matlab
45   h=s(1,1);

46

47   %% Number of columns in Reward Matrix
48   w=s(1,2);

49

50   %% CHECKING FOR SADDLE POINT &&

51

52   %% Finds equilibrium
53   P1m=[];
54   P2m=[];
55   for i=1:h
56       P1=min(R(i,1:w));
57       P1m=[P1m;P1];
58   end
59   mp1=max(P1m);
60   for j=1:w
61       P2=max(R(1:h,j));
62       P2m=[P2m;P2];
63   end
64   mp2=min(P2m);

65

66       Player1Strategy=[];
67       for i=1:h
68           if P1m(i,1)==mp1
69               P1s=1;
70           else
71               P1s=0;
72           end
73           Player1Strategy=[Player1Strategy;P1s];
74       end
75       Player2Strategy=[];
76       for i=1:w
77           if P2m(i,1)==mp2
78               P2s=1;
79           else
80               P2s=0;
81           end
82           Player2Strategy=[Player2Strategy;P2s];
83       end

84

85   if mp1==mp2 & sum(Player1Strategy)==1 & sum(Player2Strategy)==1
86       Player1Strategy=Player1Strategy;
87       Player2Strategy=Player2Strategy;
88   else
```

```
89
90  %% Linear Programming Algorithm to compute P1's optimal Strategy %%
91
92  Ar=-1*R';
93  A=[ones(w,1),Ar];
94  f=[-1, zeros(1,h)];
95  b=zeros(w,1);
96  Aeq=[0,ones(1,h)];
97  beq=1;
98  lb=[-inf;zeros(h,1)];
99  [x,fval,exitflag,output,lambda] = linprog(f,A,b,Aeq,beq,lb);
100 Player1Strategy=x(2:h+1,1);
101
102 %% Linear Programming Algorithm to compute P2's optimal Strategy %%
103
104 A2=[-ones(h,1),R];
105 f2=[1,zeros(1,w)];
106 b2=zeros(h,1);
107 Aeq2=[0,ones(1,w)];
108 beq2=1;
109 lb2=[-inf;zeros(w,1)];
110 [x2,fval,exitflag,output,lambda] = linprog(f2,A2,b2,Aeq2,beq2,lb2);
111 Player2Strategy=x2(2:w+1,1);
112 end
```

```
1   function [StrategyMat] = RunGame(R,RT,RT2)
2
3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{
5
6   Function:
7
8       [StrategyMat] = RunGame(R,RT,RT2)
9
10  Author:
11
12      Jeremy D. Jordan, Capt, USAF
13
14  Description:
15
16      Calculates the strategies of the players of a two-player
17      game.
18
19  Inputs:
```

116

```
20
21      RT: Risk tolerance of player 1
22      RT2: Risk tolerance of player 2
23
24  Outputs:
25
26      StrategyMat: Gives the strategies of the two players, the input
27      risk tolerances, and the value of the game
28
29  %}
30  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
31
32  format short
33
34  %% Size of original Reward Matrix
35  s=size(R);
36  %% Number of rows in original Reward Matrix
37  h=s(1,1);
38  %% Number of columns in original Reward Matrix
39  w=s(1,2);
40
41  %% Changes risk tolerance for expected case
42  if RT=='inf'
43      RT=0;
44  else
45  end
46
47  if RT2=='inf'
48      RT2=0;
49  else
50  end
51
52  %% Transforms original reward matrix using risk tolerance
53  [transR]=Transform(R,RT);
54
55  %% Passes transformed R into Optimal Strategy to get optimal
56  %% strategy of player 1
57  [Player1Strategy,Player2Strategy] = OptimalStrategy(transR);
58
59  %% Extracts optimal strategy of player 2 based on risk tolerance
60  [Player2Strategy]=P2Strategy(R,RT2);
61
62  % Computes the value of the game
```

```
63  ValueoftheGame=Player1Strategy'*R*Player2Strategy;
64  StrategyMat=[Player1Strategy;Player2Strategy;RT;RT2;ValueoftheGame];
```

```
1   function [transR]=Transform(R,RT)

3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{

6   Function:

8       [transR]=Transform(R,RT)

10  Author:

12      Jeremy D. Jordan, Capt, USAF

14  Description:

16      Transforms reward matrix to account for risk strategies of the
17      players of the game.

19  Inputs:

21      R: Original reward matrix

23      RT: Risk tolerance of the player

25  Output:

27      transR: The transformed reward matrix after accounting for
28      the risk strategies of the players

30  %}
31  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

33  %% Size of original Reward Matrix
34  s=size(R);
35  %% Number of rows in original Reward Matrix
36  h=s(1,1);
37  %% Number of columns in original Reward Matrix
38  w=s(1,2);

40  if RT == 0
41      transR=R;
```

118

```
42  else
43  transR=[];
44  for i=1:h
45      rjmat=[];
46      for j=1:w
47      %%Standardized Exponential
48          rj=(1-(exp(-(R(i,j)-min(min(R)))/RT)))/(1-(exp(-((max(max(R))) ...
49          -min(min(R)))/RT)));
50          rjmat=[rjmat,rj];
51      end
52  transR=[transR;rjmat];
53  end
54  end
```

```
1   function [Player2Strategy]=P2Strategy(R,RT2)
2
3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{
5
6   Function:
7
8       [Player2Strategy]=P2Strategy(R,RT2)
9
10  Author:
11
12      Jeremy D. Jordan, Capt, USAF
13
14  Description:
15
16      Gives strategy for player two based on risk tolerance
17
18  Inputs:
19
20      R: Original reward matrix
21
22      RT2: Risk tolerance of player 2
23
24  Output:
25
26      Player2Strategy: Strategy of player 2
27  %}
28  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
29
30  % Transforms the reward matrix to accomodate for player 2's risk tolerance
```

```
31   [transR]=Transform(R,RT2);

32

33   % Passes transformed reward matrix to extract optimal strategy for player 2

34   [Player1Strategy,Player2Strategy] = OptimalStrategy(transR);

35
```

```
1    function [Comparison]=natureinteractionplot(R,dist)

2

3    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

4    %{

5

6    Function:

7

8        [Comparison]=natureinteractionplot(R,dist)

9

10   Author:

11

12       Jeremy D. Jordan, Capt, USAF

13

14   Description:

15

16       Generates response surface of the risk strategy of player one and

17       the distribution of nature

18

19   Inputs:

20

21       R: The original reward matrix the strategies are calculated from

22       dist: Probability distribution of nature (player two)

23

24   Outputs:

25

26       Response surface plot of risk strategy of player one

27

28   %}

29   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

30

31   %% Size of original Reward Matrix

32   s=size(R);

33   %% Number of rows in original Reward Matrix

34   h=s(1,1);

35   %% Number of columns in original Reward Matrix

36   w=s(1,2);

37

38   valMM=[];
```

```matlab
39  RT2=0;
40  %RTM=[-50 -24.75 -1 1 24.75 50];
41  RTM=[-1:-1:-50,50:-1:1];
42  %RTM=[-1 -2 -5 -10 -20 -50 50 20 10 5 2 1];
43  for j=1:length(RTM)
44      RT=RTM(1,j);
45  [StrategyMat] = RunGame(R,RT,RT2);
46  p1strat=StrategyMat(1:h,1);
47
48  valM=[];
49  for i=1:w
50  p2truth=R(1:h,i);
51  val=p1strat'*p2truth;
52  valM=[valM;val];
53  end
54  valMM=[valMM,valM];
55  end
56
57
58  mnMat=[];
59  for i=1:length(RTM)
60      mn=dist*valMM(1:w,i);
61      mnMat=[mnMat;mn];
62  end
63  mean=mnMat;
64
65  varM=[];
66  for i=1:length(RTM)
67      vvM=[];
68      for j=1:w
69          vv=valMM(j,i)-mnMat(i,1);
70          vvM=[vvM;vv];
71      end
72      vvsqr=vvM.^2;
73      vvv=sum(vvsqr)/(w-1);
74      varM=[varM;vvv];
75  end
76  stdev=sqrt(varM);
77  variance=varM;
78
79  % varMat=[];
80  % for i=1:length(RTM)
81  %     vari=std(valMM(1:w,i));
82  %     varMat=[varMat;vari];
```

```matlab
83  % end

84

85  RPmean=mean(1:(length(RTM)/2));
86  RAmean=mean(((length(RTM)/2)+1):end);
87  RPvariance=variance(1:(length(RTM)/2));
88  RAvariance=variance(((length(RTM)/2)+1):end);

89

90  Comparison=[RTM;valMM;mean';stdev'];

91

92  subplot(2,2,1)
93  plot(RTM(1:length(RTM)/2),RPmean)
94  set(gca,'xtick',[-50 -1])
95  xlabel('Expected Case              {\rho}1              Risk Prone')
96  ylabel('y = Value of Game to Player 1')
97  axis([-51 0 min(mean)-((max(mean)-min(mean))/8)...
98      max(mean)+((max(mean)-min(mean))/8)])
99  title('Mean Value of Risk Prone Strategy')

100

101  subplot(2,2,2)
102  plot(RTM(1:length(RTM)/2),RPvariance,'r')
103  set(gca,'xtick',[-50 -1])
104  xlabel('Expected Case              {\rho}1              Risk Prone')
105  ylabel('y = Value of Game to Player 1')
106  axis([-51 0 0 ...
107      max(variance)+((max(variance)-min(variance))/8)])
108  title('Variance of Risk Prone Strategy')

109

110  subplot(2,2,3)
111  plot(RTM(((length(RTM)/2)+1):end),RAmean)
112  set(gca,'xtick',[1 50])
113  xlabel('Risk Averse              {\rho}1              Expected Case')
114  ylabel('y = Value of Game to Player 1')
115  axis([0 51 min(mean)-((max(mean)-min(mean))/8)...
116      max(mean)+((max(mean)-min(mean))/8)])
117  title('Mean Value of Risk Averse Strategy')

118

119  subplot(2,2,4)
120  plot(RTM(((length(RTM)/2)+1):end),RAvariance,'r')
121  set(gca,'xtick',[1 50])
122  xlabel('Risk Averse              {\rho}1              Expected Case')
123  ylabel('y = Value of Game to Player 1')
124  axis([0 51 0 ...
125      max(variance)+((max(variance)-min(variance))/8)])
126  title('Variance of Risk Averse Strategy')
```

```
1   function [y] = design(R,rn,ra)

2

3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

4   %{

5

6   Function:

7

8       [y] = design(R,rn,ra)

9

10  Author:

11

12      Jeremy D. Jordan, Capt, USAF

13

14  Description:

15

16      Calculates response for the designed experiment in the two

17      player game using different high and low levels.

18

19  Inputs:

20

21      R: The original reward matrix the strategies are calculated from

22      rn: Risk neutral risk parameter rho for the reward matrix

23      ra: Extreme risk averse parameter rho for the reward matrix

24

25  Outputs:

26

27      Response for the designed experiment

28  %}

29  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

30

31  rp=-ra;

32  center=(ra+rn)/2;

33

34  ymat=[];

35  for i=1:20

36      RT=[-rn;rp;-rn;rp;-rn;rp;-rn;rp;ra;rn;ra;rn;ra;rn;ra;rn;-center;-center;center;center];

37      RT2=[ra;ra;rn;rn;-rn;-rn;rp;rp;ra;ra;rn;rn;-rn;-rn;rp;rp;center;-center;center;-center];

38      RT=RT(i,1);

39      RT2=RT2(i,1);

40      [StrategyMat] = RunGame(R,RT,RT2);

41      ymat=[ymat;StrategyMat(length(StrategyMat))];

42  end

43  y=ymat;
```

```matlab
1   function interactionplot(R)
2
3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{
5
6   Function:
7
8       interactionplot(R)
9
10  Author:
11
12      Jeremy D. Jordan, Capt, USAF
13
14  Description:
15
16      Generates the response surface of the risk strategies of the players
17      using design of experiments
18
19  Inputs:
20
21      R: The original reward matrix the strategies are calculated from
22
23  Outputs:
24
25      Response surface of the risk strategies of the players of the game
26
27  %}
28  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
29
30  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
31  %% This script generates interaction plots for my design
32  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
33
34  %% Gets response for Reward Matrix
35  [y] = design(R);
36
37  %% Generates Plots for Design Region 1
38  subplot(2,2,1)
39  pneg1=[-1 1];
40  yneg1=[y(1),y(3)];
41  plot(pneg1,yneg1,'r--');
42  hold on
43  p1=[-1 1];
```

```
44 | y1=[y(2),y(4)];
45 | plot(p1,y1,'-');
46 | legend('{\rho}1 = Expected Case','{\rho}1 = Risk Prone')
47 | xlabel('Risk Averse                  {\rho}2              Expected Case')
48 | ylabel('y = Value of Game to Player 1')
49 | title('{\rho}1 {\rho}2 Interation Plot in Design Region 1')
50 | set(gca,'xtick',[-1 1])
51 | axis([-1.5 1.5 min(y)-((max(y)-min(y))/8) max(y)+((max(y)-min(y))/8)])
52 |
53 | %% Generates Plots for Design Region 2
54 | subplot(2,2,2)
55 | pneg1=[-1 1];
56 | yneg1=[y(5),y(7)];
57 | plot(pneg1,yneg1,'r--');
58 | hold on
59 | p1=[-1 1];
60 | y1=[y(6),y(8)];
61 | plot(p1,y1,'-');
62 | legend('{\rho}1 = Expected Case','{\rho}1 = Risk Prone')
63 | xlabel('Expected Case              {\rho}2              Risk Prone')
64 | ylabel('y = Value of Game to Player 1')
65 | title('{\rho}1 {\rho}2 Interation Plot in Design Region 2')
66 | set(gca,'xtick',[-1 1])
67 | axis([-1.5 1.5 min(y)-((max(y)-min(y))/8) max(y)+((max(y)-min(y))/8)])
68 |
69 |
70 | %% Generates Plots for Design Region 3
71 | subplot(2,2,3)
72 | pneg1=[-1 1];
73 | yneg1=[y(9),y(11)];
74 | plot(pneg1,yneg1,'g-');
75 | hold on
76 | p1=[-1 1];
77 | y1=[y(10),y(12)];
78 | plot(p1,y1,'r--');
79 | legend('{\rho}1 = Risk Averse','{\rho}1 = Expected Case')
80 | xlabel('Risk Averse                  {\rho}2              Expected Case')
81 | ylabel('y = Value of Game to Player 1')
82 | title('{\rho}1 {\rho}2 Interation Plot in Design Region 3')
83 | set(gca,'xtick',[-1 1])
84 | axis([-1.5 1.5 min(y)-((max(y)-min(y))/8) max(y)+((max(y)-min(y))/8)])
85 |
86 |
87 | %% Generates Plots for Design Region 4
```

```
88  subplot(2,2,4)
89  pneg1=[-1 1];
90  yneg1=[y(13),y(15)];
91  plot(pneg1,yneg1,'g-');
92  hold on
93  p1=[-1 1];
94  y1=[y(14),y(16)];
95  plot(p1,y1,'r--');
96  legend('{\rho}1 = Risk Averse','{\rho}1 = Expected Case')
97  xlabel('Expected Case              {\rho}2              Risk Prone')
98  ylabel('y = Value of Game to Player 1')
99  title('{\rho}1 {\rho}2 Interation Plot in Design Region 4')
100 set(gca,'xtick',[-1 1])
101 axis([-1.5 1.5 min(y)-((max(y)-min(y))/8) max(y)+((max(y)-min(y))/8)])
```

```
1   function contourplot(R)
2
3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{
5
6   Function:
7
8       contourplot(R)
9
10  Author:
11
12      Jeremy D. Jordan, Capt, USAF
13
14  Description:
15
16      Displays contour plot of the risk strategies of the players.
17
18  Inputs:
19
20      R: The original reward matrix
21
22  Outputs:
23
24      Contour plot
25
26  %}
27  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
28
29  %Original Units
```

```matlab
30  Xorig=[1,-50,0.5,-25;1,-0.5,0.5,-0.25;1,-50,50,-2500;1,-0.5,50,-25;1,-50,-50,2500;...
31  1,-0.5,-50,25;1,-50,-0.5,25;1,-0.5,-0.5,0.25;1,0.5,0.5,0.25;1,50,0.5,25;1,0.5,50,25;...
32  1,50,50,2500;1,0.5,-50,-25;1,50,-50,-2500;1,0.5,-0.5,-0.25;1,50,-0.5,-25];
33  %Coded Units
34  %Xorig=[1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1;1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1;...
35  1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1;1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1];
36
37  [y] = design(R);
38  yorig=y;
39
40  X=Xorig(1:4,1:4);
41  y=yorig(1:4,1);
42  B=inv(X'*X)*X'*y;
43  u=-50:.5:-.5;
44  v=.5:.5:50;
45  [P1,P2]=meshgrid(u,v);
46  Z=B(1,1)+(B(2,1).*P1)+(B(3,1)*P2)+(B(4,1).*P1.*P2);
47  cs = contour(P1,P2,Z);
48  clabel(cs)
49
50  hold on
51
52  X=Xorig(5:8,1:4);
53  y=yorig(5:8,1);
54  B=inv(X'*X)*X'*y;
55  u=-50:.5:-.5;
56  [P1,P2]=meshgrid(u,u);
57  Z=B(1,1)+(B(2,1).*P1)+(B(3,1)*P2)+(B(4,1).*P1.*P2);
58  cs = contour(P1,P2,Z);
59  clabel(cs)
60
61  hold on
62
63  X=Xorig(9:12,1:4);
64  y=yorig(9:12,1);
65  B=inv(X'*X)*X'*y;
66  u=.5:.5:50;
67  [P1,P2]=meshgrid(u,u);
68  Z=B(1,1)+(B(2,1).*P1)+(B(3,1)*P2)+(B(4,1).*P1.*P2);
69  cs = contour(P1,P2,Z);
70  clabel(cs)
71
72  hold on
73
```

```
74  X=Xorig(13:16,1:4);
75  y=yorig(13:16,1);
76  B=inv(X'*X)*X'*y;
77  v=-50:.5:-.5;
78  u=.5:.5:50;
79  [P1,P2]=meshgrid(u,v);
80  Z=B(1,1)+(B(2,1).*P1)+(B(3,1)*P2)+(B(4,1).*P1.*P2);
81  cs = contour(P1,P2,Z);
82  clabel(cs)
83
84  axis([-50 50 -50 50]);
85
86  xlabel('Player 1')
87  ylabel('Player 2')
88  zlabel('Value of Game to Player 1')
89  title('Response Surface of Game')
```

```
1
2   function responsesurface(R)
3
4   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5   %{
6
7   Function:
8
9       responsesurface(R)
10
11  Author:
12
13      Jeremy D. Jordan, Capt, USAF
14
15  Description:
16
17      Gives response surface in 3-d, countour, and two-way interaction
18      plots.
19
20  Inputs:
21
22      R: The original reward matrix the strategies are calculated from
23
24  Outputs:
25
26      3-d response surface
27      Contour plot of response surface
```

128

```matlab
28       2-way interaction plot of response surface
29
30   %}
31   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
32
33   low=1;
34   high=50;
35   %Original Units
36   Xorig=[1,-high,low,-high*low;1,-low,low,-low*low;1,-high,high,-high*high;...
37   1,-low,high,-low*high;1,-high,-high,-high*-high;1,-low,-high,-low*-high;...
38   1,-high,-low,-high*-low;1,-low,-low,-low*-low;1,low,low,low*low;1,high,low,...
39   high*low;1,low,high,low*high;1,high,high,high*high;1,low,-high,low*-high;1,...
40   high,-high,high*-high;1,low,-low,low*-low;1,high,-low,high*-low];
41   %Coded Units
42   %Xorig=[1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1;1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;...
43   1,1,1,1;1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;1,1,1,1;1,-1,-1,1;1,1,-1,-1;1,-1,1,-1;...
44   1,1,1,1];
45   s=size(Xorig);
46   %columns in X
47   w=s(1,2);
48   %rows in X
49   h=s(1,1);
50
51   [y] = design(R);
52   %% Design Region 1
53   yorig=y;
54   figure(1)
55   X=Xorig(1:4,1:4);
56   y=yorig(1:4,1);
57   B=inv(X'*X)*X'*y;
58   u=-high:.5:-low;
59   v=low:.5:high;
60   [P1,P2]=meshgrid(u,v);
61   Z=B(1,1)+(B(2,1).*P1)+(B(3,1).*P2)+(B(4,1).*P1.*P2);
62   surf(P1,P2,Z)
63
64   hold on
65   %% Design Region 2
66   X=Xorig(5:8,1:4);
67   y=yorig(5:8,1);
68   B=inv(X'*X)*X'*y;
69   u=-high:.5:-low;
70   [P1,P2]=meshgrid(u,u);
71   Z=B(1,1)+(B(2,1).*P1)+(B(3,1).*P2)+(B(4,1).*P1.*P2);
```

129

```matlab
72   surf(P1,P2,Z)
73
74   hold on
75   %% Design Region 3
76   X=Xorig(9:12,1:4);
77   y=yorig(9:12,1);
78   B=inv(X'*X)*X'*y;
79   u=low:.5:high;
80   [P1,P2]=meshgrid(u,u);
81   Z=B(1,1)+(B(2,1).*P1)+(B(3,1).*P2)+(B(4,1).*P1.*P2);
82   surf(P1,P2,Z)
83
84   hold on
85   %% Design Region 4
86   X=Xorig(13:16,1:4);
87   y=yorig(13:16,1);
88   B=inv(X'*X)*X'*y;
89   v=-high:.5:-low;
90   u=low:.5:high;
91   [P1,P2]=meshgrid(u,v);
92   Z=B(1,1)+(B(2,1).*P1)+(B(3,1).*P2)+(B(4,1).*P1.*P2);
93   surf(P1,P2,Z)
94
95   xlabel('Player 1')
96   ylabel('Player 2')
97   zlabel('Value of Game to Player 1')
98   title('Response Surface of Game')
99
100  figure(2)
101  interactionplot(R)
102  figure(3)
103  contourplot(R)
```

```matlab
1    function explore(R)
2
3    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4    %{
5
6    Function:
7
8        explore(R)
9
10   Author:
11
```

```
12        Jeremy D. Jordan, Capt, USAF

13

14  Description:

15

16        Explores the true response surface of the two player game

17

18  Inputs:

19

20        R: The original reward matrix the strategies are calculated from

21

22  Outputs:

23

24        True response surface of the two-player game. This is complimentary

25        to interactionplot.m which explores the response surface through DOE

26

27  %}
28  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

29

30  low=1;
31  high=50;
32  int=1;
33  Risklevel=1;

34

35  %% Creates Response Surface of Design Region 1

36

37  y=[];
38  RTol2=low:int:high;
39          RT=-Risklevel;
40      for j=1:length(RTol2)
41          RT2=RTol2(j);
42          [StrategyMat] = RunGame(R,RT,RT2);
43          GameVal(j)=StrategyMat(end,1);
44      end
45      y=GameVal';

46

47      for i=1:length(RTol2)
48          xmat(i,1:2)=[RT,RTol2(i)];
49      end
50      X=xmat;
51  subplot(2,2,1)
52  plot(X(:,2),y)
53  hold on
54  y1=[];
55  RTol2=low:int:high;
```

```
56        RT=0;
57     for j=1:length(RTol2)
58        RT2=RTol2(j);
59        [StrategyMat] = RunGame(R,RT,RT2);
60        GameVal(j)=StrategyMat(end,1);
61     end
62     y1=GameVal';;
63
64     for i=1:length(RTol2)
65        xmat(i,1:2)=[RT,RTol2(i)];
66     end
67     X=xmat;
68  plot(X(:,2),y1,'r')
69  legend('{\rho}1 = Risk Prone' ,'{\rho}1 = Expected Case')
70  xlabel('Risk Averse                {\rho}2            Expected Case')
71  ylabel('y = Value of Game to Player 1')
72  title('{\rho}1 {\rho}2 Interation Plot in Design Region 1')
73
74  axis([.5 50.5 (min(min(y,y1)))-((max(max(y,y1)))-(min(min(y,y1))))...
75   (max(max(y,y1)))+(max(max(y,y1))-(min(min(y,y1))))])
76
77  %% Creates Response Surface of Design Region 2
78
79  y=[];
80  RTol2=-high:int:-low;
81        RT=-Risklevel;
82     for j=1:length(RTol2)
83        RT2=RTol2(j);
84        [StrategyMat] = RunGame(R,RT,RT2);
85        GameVal(j)=StrategyMat(end,1);
86     end
87     y=GameVal';
88
89     for i=1:length(RTol2)
90        xmat(i,1:2)=[RT,RTol2(i)];
91     end
92     X=xmat;
93  subplot(2,2,2)
94  plot(X(:,2),y)
95  hold on
96  y1=[];
97  RTol2=-high:int:-low;
98        RT=0;
99     for j=1:length(RTol2)
```

```matlab
100          RT2=RTol2(j);
101          [StrategyMat] = RunGame(R,RT,RT2);
102          GameVal(j)=StrategyMat(end,1);
103      end
104      y1=GameVal';
105
106      for i=1:length(RTol2)
107          xmat(i,1:2)=[RT,RTol2(i)];
108      end
109      X=xmat;
110  plot(X(:,2),y1,'r')
111  legend('{\rho}1 = Risk Prone','{\rho}1 = Expected Case')
112  xlabel('Expected Case              {\rho}2              Risk Prone')
113  ylabel('y = Value of Game to Player 1')
114  title('{\rho}1 {\rho}2 Interation Plot in Design Region 2')
115
116  axis([-50.5 -.5 (min(min(y,y1)))-((max(max(y,y1)))-(min(min(y,y1))))...
117   (max(max(y,y1)))+(max(max(y,y1))-(min(min(y,y1))))])
118
119  %% Creates Response Surface of Design Region 3
120
121  y=[];
122  RTol2=low:int:high;
123          RT=Risklevel;
124      for j=1:length(RTol2)
125          RT2=RTol2(j);
126          [StrategyMat] = RunGame(R,RT,RT2);
127          GameVal(j)=StrategyMat(end,1);
128      end
129      y=GameVal';
130
131      for i=1:length(RTol2)
132          xmat(i,1:2)=[RT,RTol2(i)];
133      end
134      X=xmat;
135  subplot(2,2,3)
136  plot(X(:,2),y)
137  hold on
138  y1=[];
139  RTol2=low:int:high;
140          RT=0;
141      for j=1:length(RTol2)
142          RT2=RTol2(j);
143          [StrategyMat] = RunGame(R,RT,RT2);
```

```matlab
144          GameVal(j)=StrategyMat(end,1);
145      end
146      y1=GameVal';;
147
148      for i=1:length(RTol2)
149          xmat(i,1:2)=[RT,RTol2(i)];
150      end
151      X=xmat;
152  plot(X(:,2),y1,'r')
153  legend('{\rho}1 = Risk Averse' ,'{\rho}1 = Expected Case')
154  xlabel('Risk Averse                {\rho}2             Expected Case')
155  ylabel('y = Value of Game to Player 1')
156  title('{\rho}1 {\rho}2 Interation Plot in Design Region 3')
157
158  axis([.5 50.5 (min(min(y,y1)))-((max(max(y,y1)))-(min(min(y,y1))))...
159    (max(max(y,y1)))+(max(max(y,y1))-(min(min(y,y1))))])
160
161  %% Creates Response Surface in Design region 4
162
163  y=[];
164  RTol2=-high:int:-low;
165          RT=Risklevel;
166      for j=1:length(RTol2)
167          RT2=RTol2(j);
168          [StrategyMat] = RunGame(R,RT,RT2);
169          GameVal(j)=StrategyMat(end,1);
170      end
171      y=GameVal';
172
173      for i=1:length(RTol2)
174          xmat(i,1:2)=[RT,RTol2(i)];
175      end
176      X=xmat;
177  subplot(2,2,4)
178  plot(X(:,2),y)
179  hold on
180  y1=[];
181  RTol2=-high:int:-low;
182          RT=0;
183      for j=1:length(RTol2)
184          RT2=RTol2(j);
185          [StrategyMat] = RunGame(R,RT,RT2);
186          GameVal(j)=StrategyMat(end,1);
187      end
```

```
188    y1=GameVal';
189
190    for i=1:length(RTol2)
191        xmat(i,1:2)=[RT,RTol2(i)];
192    end
193    X=xmat;
194  plot(X(:,2),y1,'r')
195  legend('{\rho}1 = Risk Averse','{\rho}1 = Expected Case')
196  xlabel('Expected Case            {\rho}2            Risk Prone')
197  ylabel('y = Value of Game to Player 1')
198  title('{\rho}1 {\rho}2 Interation Plot in Design Region 4')
199
200  axis([-50.5 -.5 (min(min(y,y1)))-((max(max(y,y1)))-(min(min(y,y1))))...
201    (max(max(y,y1)))+(max(max(y,y1)))-(min(min(y,y1))))])
```

```
1   function exploreexactsurface(R)
2
3   %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4   %{
5
6   Function:
7
8       exploreexactsurface(R)
9
10  Author:
11
12      Jeremy D. Jordan, Capt, USAF
13
14  Description:
15
16      Gives exact response surface in 3-dimensional graphs
17
18  Inputs:
19
20      R: The original reward matrix the strategies are calculated from
21
22  Outputs:
23
24      3-d graph of response surface
25
26  %}
27  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
28
29  low=1;
```

```matlab
30  high=50;
31  int=1;
32
33  %%Creates Response Surface of Design Region 1
34  figure(4)
35  y=[];
36
37  RTol1=-low:-int:-high;
38  RTol2=low:int:high;
39  for i=1:length(RTol1)
40      RT=RTol1(i);
41      GameVal=[];
42      for j=1:length(RTol2)
43          RT2=RTol2(j);
44          [StrategyMat] = RunGame(R,RT,RT2);
45          GameVal(j)=StrategyMat(end,1);
46      end
47      y=[y;GameVal'];
48  end
49
50  X=[];
51  for j=1:length(RTol1)
52      xmat=[];
53      for i=1:length(RTol2)
54          xmat(i,1:2)=[RTol1(j),RTol2(i)];
55      end
56      X=[X;xmat];
57  end
58  scatter3(X(:,1),X(:,2),y,'.');
59  xi=linspace(-high,-low,100);
60  yi=linspace(low,high,100)';
61  [Xi,Yi,Zi]=griddata(X(:,1),X(:,2),y,xi,yi,'v4');
62  surf(Xi,Yi,Zi)
63
64  hold on
65
66  %% Creates Response Surface of Design Region 2
67  y=[];
68  RTol1=-low:-int:-high;
69  RTol2=-low:-int:-high;
70  for i=1:length(RTol1)
71      RT=RTol1(i);
72      GameVal=[];
73      for j=1:length(RTol2)
```

```matlab
74          RT2=RTol2(j);
75          [StrategyMat] = RunGame(R,RT,RT2);
76          GameVal(j)=StrategyMat(end,1);
77      end
78      y=[y;GameVal'];
79  end
80
81  X=[];
82  for j=1:length(RTol1)
83      xmat=[];
84      for i=1:length(RTol2)
85          xmat(i,1:2)=[RTol1(j),RTol2(i)];
86      end
87      X=[X;xmat];
88  end
89  scatter3(X(:,1),X(:,2),y,'.');
90  xi=linspace(-high,-low,100);
91  yi=linspace(-high,-low,100)';
92  [Xi,Yi,Zi]=griddata(X(:,1),X(:,2),y,xi,yi,'v4');
93  surf(Xi,Yi,Zi)
94
95
96  hold on
97
98  %%Creates Response Surface of Design Region 3
99  y=[];
100 RTol1=low:int:high;
101 RTol2=low:int:high;
102 for i=1:length(RTol1)
103     RT=RTol1(i);
104     GameVal=[];
105     for j=1:length(RTol2)
106         RT2=RTol2(j);
107         [StrategyMat] = RunGame(R,RT,RT2);
108         GameVal(j)=StrategyMat(end,1);
109     end
110     y=[y;GameVal'];
111 end
112
113 X=[];
114 for j=1:length(RTol1)
115     xmat=[];
116     for i=1:length(RTol2)
117         xmat(i,1:2)=[RTol1(j),RTol2(i)];
```

137

```matlab
118        end
119        X=[X;xmat];
120    end
121    scatter3(X(:,1),X(:,2),y,'.');
122    xi=linspace(low,high,100);
123    yi=linspace(low,high,100)';
124    [Xi,Yi,Zi]=griddata(X(:,1),X(:,2),y,xi,yi,'v4');
125    surf(Xi,Yi,Zi)
126
127    hold on
128
129    %%Creates Response Surface of Design Region 4
130    y=[];
131    RTol1=low:int:high;
132    RTol2=-low:-int:-high;
133    for i=1:length(RTol1)
134        RT=RTol1(i);
135        GameVal=[];
136        for j=1:length(RTol2)
137            RT2=RTol2(j);
138            [StrategyMat] = RunGame(R,RT,RT2);
139            GameVal(j)=StrategyMat(end,1);
140        end
141        y=[y;GameVal'];
142    end
143
144    X=[];
145    for j=1:length(RTol1)
146        xmat=[];
147        for i=1:length(RTol2)
148            xmat(i,1:2)=[RTol1(j),RTol2(i)];
149        end
150        X=[X;xmat];
151    end
152    scatter3(X(:,1),X(:,2),y,'.');
153    xi=linspace(low,high,100);
154    yi=linspace(-high,-low,100)';
155    [Xi,Yi,Zi]=griddata(X(:,1),X(:,2),y,xi,yi,'v4');
156    surf(Xi,Yi,Zi)
157    %
158    % responsesurface(R)
```

```matlab
1    function  simulation(R,N,Combat)
2
```

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%{

Function:

    simulation(R,N,Combat)

Author:

    Jeremy D. Jordan, Capt, USAF

Description:

    Simulates responses from subject matter experts and randomizes
    reward matrix

Inputs:

    R: The original reward matrix the strategies are calculated from
    N: Number of survey responses
    Combat: 1 if using a Likert scale between -5 and 5, 0 if no
        restriction on reward matrix

Outputs:

    Simulated vs true value of the game and randomized reward matrix
    if desired.

%}
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Combat = 1 or 0, 1 if this is a combat scenario
%Use for randomizing reward matrix
%[R] =

s=size(R);
h=s(1,1);
w=s(1,2);
origR=R;
RT=0;
RT2=0;

[StrategyMat,OptimalRisk,ValueofGame] = RunGameOld(R,RT,RT2);
TrueStrategyMat=StrategyMat;
```

```matlab
47   TrueOptimalRisk=OptimalRisk;

48

49

50   OptimalRiskMat=[];
51   DifferenceMat=[];
52   AverageOptimalRiskMat=[];

53

54

55   for i=1:N

56

57   Rmat=[];
58   for i=1:h
59       rrmat=[];
60       for j=1:w
61           %% USE FOR RANDOMIZING REWARD MATRIX %%
62           %rr = random('Normal',origR(i,j),1,1,1);
63           rr=randsrc(1,1,[origR(i,j),origR(i,j)-1,origR(i,j)+1;.5,.25,.25]);
64           rrmat=[rrmat,rr];
65       end
66       Rmat=[Rmat;rrmat];
67   end
68   R=Rmat;
69   R = roundn(Rmat,-1);

70

71   %% Use to keep values of R within 5 %%
72   if Combat==1
73       for i=1:h
74           for j=1:w
75               if R(i,j)>5
76                   R(i,j)=5;
77               elseif R(i,j)<-5
78                   R(i,j)=-5;
79               end
80           end
81       end
82   else
83   end

84

85   [StrategyMat,OptimalRisk,ValueofGame] = RunGameOld(R,RT,RT2);
86   OptimalRiskMat=[OptimalRiskMat;OptimalRisk(1,1)];

87

88   AverageOptimalRisk=mean(OptimalRiskMat);
89   AverageOptimalRiskMat=[AverageOptimalRiskMat;AverageOptimalRisk];

90
```

140

```matlab
91  Difference=TrueOptimalRisk(1,1) - AverageOptimalRisk;
92  DifferenceMat=[DifferenceMat;Difference];
93
94  end
95
96  n=[1:N];
97  TrueOptimalRiskMat=[];
98  for i=1:N
99      TO=TrueOptimalRisk(1,1);
100     TrueOptimalRiskMat=[TrueOptimalRiskMat,TO];
101 end
102
103
104 figure(1)
105 plot(n,DifferenceMat)
106 xlabel('N')
107 ylabel('Value of Game \pi')
108 title('Difference between true value and simulated value')
109
110 figure(2)
111 plot(n,AverageOptimalRiskMat)
112 xlabel('N')
113 ylabel('Value of Game \pi')
114 title('Value of Game as N increases')
115 hold on
116 plot(n,TrueOptimalRiskMat','r')
117 legend('Simulated Value of Game','Red=True Value of Game')
118
```

141

Bibliography

1. Arrow, Kenneth. (1957). Decision Theory and Operations Research. *Operations Research*, **5**, (6) 765-774.

2. Athans, Michael. (1987). Command and Control (C2) Theory: A Challenge to Control Science. *IEEE Transactions on Automatic Control*, **A-C 32** (4), Online.

3. Berkovitz, L.D. and Dresher, M. (1959). A Game-Theory Analysis of Tactical Air-War. *Operations Research*, **7** (5) 599-620.

4. Bier, Vicki M. and Abhichandani, Vinod. (2002). Optimal Allocation of Resources for Defense of Simple Series and Parallel Systems from Determined Adversaries. *Proceedings of the Engineering Foundation Conference on Risk Based Decision Making in Water Resources X*, American Society of Civil Engineers, Santa Barbara, CA.

5. Caywood, T.E. and Thomas, C.J. (1955). Applications of Game Theory in Fighter versus Bomber Combat. *Journal of the Operations Research Society of America*, **3**, (4) 402-411.

6. Charnes, A., Kirby, M., and Raike, W. (1968). Chance-Constrained Games with Partially Controllable Strategies. *Operations Research*, **16**, (1) 142-149.

7. Cooman, G. and Zaffalon, M. (2004). Updating Beliefs with Incomplete Observations. *Artificial Intelligence*, **159**, 75-125.

8. Cruz, J.B., Simaan, M.A., Gacic, A., Jiang, H., Letellier, B. and Li, M. (2000). Modeling and Control of Military Operations Against Adversarial Control. *Proceedings of the 39th IEEE Conference on Decision and Control*, **3** 2581-2586.

9. Cruz, J.B., Simaan, M.A., Gacic, A., Jiang, H., Letellier, B., Li, M. and Liu, Y. (2001). Game-Theoretic Modeling and Control of a Military Air Operation. *IEEE Transactions on Aerospace and Electronic Systems*, **37**, (4) 1393-1405.

10. Harris, Bernard (2004). *Mathematical Methods in Combatting Terrorism. Risk Analysis*, **24** (4), 985-988.

11. Kirkwood, C. W. (1997). Strategic Decision Making. *Wadsworth Publishing Company*, **6**, pp 64-68.

12. Krichman, M., Ghose, D., Speyer, J.L., and Shamma, J.S. (2001). Theater Level Campaign Resource Allocation. *Proceedings of the American Control Conference*, **6** 4716-4721.

13. Lipovetsky (2007). Antagonistic and Bargaining Games in Optimal Marketing Decisions. *International Journal of Mathematical Education in Science and Technology*, **38** (1), 103-113.

14. McEneaney, William M. and Singh, Rajdeep (2004). Unmanned Vehicle Decision Making under Imperfect Information in an Adversarial Environment. *Proceedings of 2004 AIAA Guidance, Navigation, and Control Conference and Exhibit*, Available Online.

15. Nash, J. (1951) "Non-Cooperative Games". *Annals of Mathematics*, **54** (2), 286-295.

16. O'Neill, B. (1987). Nonmetric Test of the Minimax Theory of Two-Person Zero-sum Games. *Economic Sciences*, **84**, 2106-2109.

17. Ozdemirel, N.E. and Kandiller L. (2006). Semi-Dynamic Modeling of Heterogenous Land Combat. *Journal of the Operational Research Society*, **57**, 38-51.

18. Perry, W.L. and Moffat, J. (1997). Measuring the Effects of Knowledge in Military Campaigns *Journal of the Operations Research Society of America*, **48**, 965-972.

19. Pugh, George E. and Mayberry, John P. (1973). Theory of Measures of Effectiveness for General-Purpose Military Forces: Part I. A Zero-Sum Payoff Appropriate for Evaluating Combat Strategies. *Operations Research*, **21**, 867-885.

20. Robinson, Julia. (1951). An Iterative Method of Solving a Game *Annals of Mathematics*, **54**, (2) 296-301.

21. Sandler, Todd and Arce, Daniel (2003). Terroism and Game Theory. *Simulation Gaming*, **3434**, (3) 319-337.

22. Sandoy, Marlene and Aven, Terje. (2006). Real Time Updating of Risk Assessments During a Drilling Operation - Alternative Bayesian Approaches. *International Journal of Reliability, Quality and Safety Engineering*, **13** (1), 85-95.

23. Shor, Mikhael. Dictionary of Game Theory Terms, Game Theory .net, http://www.gametheory.net/dictionary/ Web accessed: October 15, 2006.

24. Solan, E. and Yariv, L. (2004). Games with Espionage *Games and Economic Behavior*, **47**, 172-199.

25. Stennek, Johan. (2000). The Survival Value of Assuming Others to be Rational. *International Journal on Game Theory*, **29**, 147-163.

26. Sujit, P.B. and Ghose, D. (2005). Search by UAVs with Flight Time Constraints using Game Theoretic Models. *Proceedings of the AIAA Guidance, Navigation, and Conference and Exhibit*, **AIAA 2005-6241**.

27. Thomas, Clayton J. and Deemer, Walter L. (1957). The Role of Operational Gaming in Operations Research. *Operations Research*, **5** (1) 1-27.

28. Von Neumann, J. and Morgenstern, O. (1944). Theory of Games and Economic Behavior. *Science Editions*, **3**, 1964.

29. Whittaker, G. M. (2000). Asymmetric Wargaming: Toward a Game Theoretic Perspective *In the Mitre Digest*, Available Online.

30. Winston, W. L. (2004). Operations Research: Applications and Algorithms. *Thomson Learning*, **6**.

31. Zhuang, Jun and Bier, Vicki (2006). Balancing Terrorism and Natural Disasters - Defensive Strategy with Endogenous Attacker Effort. *Submitted to Operations Research.*

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD–MM–YYYY)	2. REPORT TYPE	3. DATES COVERED (From — To)
15–03–2007	Master's Thesis	Sept 2005 — Mar 2007

4. TITLE AND SUBTITLE

Updating Optimal Decisions Using Game Theory
and Exploring Risk Behavior
Through Response Surface Methodology

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Jeremy D. Jordan, Capt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GOR/ENS/07-13

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

None

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approval for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This thesis utilizes game theory within a framework for updating optimal decisions based on new information as it becomes available. Methodology is developed that allows a decision maker to change his perceived optimal policy based on available knowledge of the opponents strategy, where the opponent is a rational decision maker or a random component nature. Utility theory is applied to account for the different risk preferences of the decision makers. Furthermore, response surface methodology is used to explore good risk strategies for the decision maker to approach each situation with. The techniques are applied to a combat scenario, a football game, and a terrorist resource allocation problem, providing a decision maker with the best possible strategy given the information available to him. The results are intuitive and exemplify the benefits of using the methods.

15. SUBJECT TERMS

Game Theory, Updating Decisions, Optimal Decision Policy Updating, Combat Modeling, Utility Theory, Risk Behavior

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Marcus Perry, Dr. Sharif Melouk
U	U	U	UU	157	19b. TELEPHONE NUMBER (include area code) (937) 255–3636, ext 4588